**PCT** WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/00, 15/12, 15/63
C12N 15/70, 15/74, 15/79

A1

(11) International Publication Number: WO 93/04166

(43) International Publication Date: 4 March 1993 (04.03.93)

(21) International Application Number: PCT/US92/06748

(22) International Filing Date: 17 August 1992 (17.08.92)

(30) Priority data:
746,032 16 August 1991 (16.08.91) US

(71) Applicant: CARNEGIE INSTITUTION OF WASHINGTON [US/US]; 1530 P Street N.W., Washington, DC 20005 (US).

(72) Inventors: MCKNIGHT, Steven, L. ; 7113 Oxford Road, Baltimore, MD 21212 (US). THOMPSON, Catherine, C. ; 1325 West 42nd Street, Baltimore, MD 21211 (US). LAMARCO, Karen, L. ; 1545 18th Street N.W. #708, Washington, DC 20036 (US).

(74) Agents: KOKULIS, Paul, N. et al.; Cushman, Darby & Cushman, 1615 L Street N.W., Washington, DC 20036 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, SE).

**Published**
*With international search report.*

(54) Title: PURINE-REGION DNA BINDING PROTEIN

(57) Abstract

The present invention relates to a GA binding protein and to DNA segments encoding subunits thereof.

*FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | | | | | |
|---|---|---|---|---|---|
| AT | Austria | FI | Finland | MN | Mongolia |
| AU | Australia | FR | France | MR | Mauritania |
| BB | Barbados | GA | Gabon | MW | Malawi |
| BE | Belgium | GB | United Kingdom | NL | Netherlands |
| BF | Burkina Faso | GN | Guinea | NO | Norway |
| BG | Bulgaria | GR | Greece | NZ | New Zealand |
| BJ | Benin | HU | Hungary | PL | Poland |
| BR | Brazil | IE | Ireland | PT | Portugal |
| CA | Canada | IT | Italy | RO | Romania |
| CF | Central African Republic | JP | Japan | RU | Russian Federation |
| CG | Congo | KP | Democratic People's Republic of Korea | SD | Sudan |
| CH | Switzerland | | | SE | Sweden |
| CI | Côte d'Ivoire | KR | Republic of Korea | SK | Slovak Republic |
| CM | Cameroon | LI | Liechtenstein | SN | Senegal |
| CS | Czechoslovakia | LK | Sri Lanka | SU | Soviet Union |
| CZ | Czech Republic | LU | Luxembourg | TD | Chad |
| DE | Germany | MC | Monaco | TG | Togo |
| DK | Denmark | MG | Madagascar | UA | Ukraine |
| ES | Spain | ML | Mali | US | United States of America |

# PURINE-REGION DNA BINDING PROTEIN

## BACKGROUND OF THE INVENTION

### Technical Field

The present invention relates, in general, to a DNA binding protein and to a DNA sequence encoding same. In particular, the invention relates to a GA binding protein and to DNA segments encoding the subunits thereof.

### Background Information

Herpes simplex virus 1 (HSV1) immediate early (IE) genes are induced at the outset of the lytic infection by a virion associated protein termed VP16 (Post et al, Cell 24, 555 (1981)). At least two classes of *cis*-regulatory elements qualify HSV IE genes for induction by VP16. The most essential VP16 *cis*-response element is characterized by the nonanucleotide sequence 5'-TAATGARAT-3' (Mackem et al, Proc Natl. Acad. Sci. U.S.A. 79, 4917 (1982); Mackem et al, J. Virol. 44, 939 (1982); Cordingley et al, Nucleic Acids Res. 11, 2347 (1983); Kristie et al, Proc. Natl. Acad. Sci. U.S.A. 81, 4065 (1984); Gaffney et al, Nucleic Acids Res. 13, 7847 (1985); Bzik et al, *ibid*. 14, 929 (1986); O'Hare et al, J. Virol. 61, 190 (1987); and Triezenberg et al, Genes Dev. 2, 730 (1988)). VP16 binds tightly to this DNA sequence in a complex with the cellular transcription factor Oct1 (Preston et al, C ll 52, 425 (1988); O'Hare et al, *ibid*. p. 435 (1988); and Gerster et al, Proc. Natl. Acad. Sci.

U.S.A. 85, 6347 (1988)). A second *cis*-regulatory element required for VP16-mediated induction of HSV IE genes consists of three imperfect repeats of the purine-rich hexanucleotide 5'-CGGAAR-3' (Triezenberg et al, Genes Dev. 2, 730 (1988) and Spector et al, *ibid.* 87, 5268 (1990)). A protein complex capable of avid interaction with the purine-rich repeats (GA repeats) has been identified in soluble preparations of rat liver nucleic (Triezenberg et al, Genes Dev. 2, 730 (1988)). This GA binding protein (GABP) consists of two separable subunits. Purified samples of either subunit do not interact with the GA repeats, yet regain potent DNA binding activity when mixed (LaMarco et al, Genes Dev. 3, 1372 (1989)).

Applicants have isolated cDNA clones encoding both subunits of GABP and have revealed that one (GABPα) is related to the Ets transforming protein, while the other (GABPβ) contains a series of 33-amino acid repeats related in sequence to a variety of proteins including Notch of Drosophila melanogaster, Lin12 and Glp1 of Caenorhabditis elegans and SW14 and SW16 of Saccharomyces cerevisiae (Wharton et al, Cell 43, 567 (1985); Greenwald, *ibid* 43, 583 (1985); Yochem et al, Nature 335, 547 (1988); Yochem et al, Cell 58, 553 (1989); Austin et al, *ibid.*, p. 565 (1989); Breeden et al, Nature 329, 651 (1987); and Andrews et al, *ibid.* 342, 830 (1989)). In addition, Applicants have demonstrated that these two protein sequence motifs, the Ets-related domain of GABPα and the 33-amino acid repeats of GABPβ, contribute the surfaces that form a multiprotein complex capable of stable and specific interaction with DNA. These findings, which form the basis of the present invention, provide insight into the problem of regulatory specificity and define, for the first time, a

discrete function for the 33-amino acid repeat motif.

## SUMMARY OF THE INVENTION

It is a general object of the invention to provide DNA segments encoding the subunits of GABP.

In one embodiment, the present invention relates to a DNA segment encoding GABPα, GABPβ1 or GABPβ2, or portion thereof.

In a further embodiment, the present invention relates to a construct comprising at least one of the above-described segments and to a host cell transformed therewith.

Further objects and advantages of the present invention will become clear from the description that follows.

## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Amino acid sequences of tryptic peptides derived from GA binding proteins. GABP (20 μg) was purified to homogeneity (inset) as described (LaMarco et al, Genes Dev. 3, 1372 (1989)) except that boiled salmon sperm DNA (20 μg/ml) was included as a non-specific competitor in the DNA affinity chromatographic step. Approximately 500 picomoles of protein was lyophilized, reduced, acetylated, and subjected to cleavage by trypsin (Boehringer Mannheim). The resulting peptides were separated by reverse-phase HPLC as described (Stone et al, Laboratory Methodology in Biochemistry, Fini et al eds (RC Press, Boca Raton, FL, 1991)). Amino acid sequence analysis was performed on a Vydac C18 column (Appli d Biosystems 477-A). The amino acid sequences derived from peaks 1-13 w re: 1, SLFDQGVIEK; 2, ?AWALEGY; 3, DEIS?VGDEGEFK; 4,

ELESLNQEDFFQR; 5, LQESLDAHEIELQDIQL?P?R; 6, DQISIVGDEGEFK; 7, MAELV; 8, YVQASQLQQMNEIVTIDQP; 9, TPLHMWASEGHA; 10, GEILWS; 11, LIEIEIDGTEK; 12, ILMANGAPFTTD; 13, TGNNGQIQL?QFLLEL?TDR.

Figure 2. Nucleotide and deduced amino acid sequences of cDNAs encoding GABP subunits. (A) Sequence of GABPα. (B) Sequences of GABPβ1 and β2. An unamplified cDNA library prepared from mouse adipocyte mRNA was screened with a mixture of degenerate oligonucleotides derived from the amino acids sequences of peptides 3, 4, 5, and 8 (Fig. 1) labeled with $^{32}P$ using polynucleotide kinase. The basic SSC protocol was used (Ausubel et al, Current Protocols in Molecular Biology (Wiley & Sons, NY), 1989). Hybridization was performed at 48°C for 16 hrs. GABPβ1 and β2 were isolated by screening a day-8.5 mouse embryo cDNA library (Lee, Mol. Endocrinol. 4, 1034 (1990)) with degenerate oligonucleotides corresponding to peptides 9 and 12 (Fig. 1). Kinased oligonucleotide probes were hybrided in 6X SSC, 1X Denhardt's, 0.05% sodium pyrophosphate, and 100 μg/ml yeast tRNA at 50°C for 14 hours. Washing conditions were 6X SSC, 0.05% sodium pyrophosphate at 55°C. A total of five clones were isolated that hybridized with both oligonucleotide probes. Four of the clones were approximately 2.6 kb and differed only slightly in the length of the 5' untranslated region; these cDNA clones encoded GABPβ1. The fifth cDNA clone was approximately 1.4 kb and differed from the other four at its 3' end; this cDNA clone encoded GABPβ2. Four additional cDNA clones corresponding to GABPβ2 were subsequently identified. DNA sequencing was by the dideoxy chain termination method (Sanger et al, Proc. Natl. Acad. Sci. USA 74, 5463 (1977)) using Sequenase (U.S. Biochemicals) under conditions

suggested by the manufacturer. The complete nucleotide sequences of GABPα, β1 and β2 were determined on both DNA strands using deleted templates or synthetic oligonucleotide primers. Deletions were made using exonuclease III (Pharmacia) under conditions specified by the manufacturer. The sequences for GABPβ1 and GABPβ2 were identical up to nucleotide 1130 except for a three nucleotide insertion (GTA) at position 828 of GABPβ1. Sequencing of four other independent isolates of GABPβ1 were identical to GABPβ2 at this site. Peptides identified by amino acid sequencing of purified GABP are underlined in the deduced amino acid sequences. The dashed lines indicate the sequence in GABPβ1 not found in GABPβ2. The sequence for β2 is shown from the point at which it diverges from β1.

Figure 3. Tissue distribution of GABP mRNAS. RNA was isolated from various rat tissues (Chingwin et al, Biochemistry 18, 5294 (1979)) and mouse L cells (Chomczynski et al, Anal. Biochem. 162, 156 (1987)). 10μg of poly A+ RNA was separated on a 1% agarose-formaldehyde gel, transferred to Nytran (Schleicher and Schuell) and hybridized with a random-primed probe prepared from GABPα (A) or GABPβ1 (B).

Figure 4. Requirement of GABPα and GABPβ1 for sequence specific DNA binding. A) In vitro translation of GABP proteins. Sense and anti-sense RNAs from GABPα and GABPβ were transcribed in vitro and translated in rabbit reticulocyte lysates (Krieg et al, Nucl. Acids Res. 12, 7057 (1984)). Plasmids with cDNAs inserted in the Eco RI and Xho I sites of Bluescript (Stratagene) were linearized with Asp 718 and transcribed with T3 RNA polymerase to generate sense strand RNA, or linearized with Bam HI and

transcribed with T7 RNA polymerase to generate anti-sense RNA. RNAs were used to program rabbit reticulocyte lysates in the presence of $^{35}$S-methionine under conditions specified by the manufacturer (Promega Biotec). Unlabeled protein was used for DNA binding experiments. The $^{35}$S-methionine labeled products were separated on a 12.5% SDS-polyacrylamide gel and visualized by fluorography; (-), anti-sense RNA; (+), sense strand RNA. Positions of molecular weight markers are indicated in kD. (B) Electrophoretic mobility shift assays with <u>in vitro</u> translated GABP proteins. Proteins were incubated in the presence of a $^{32}$P-labeled DNA fragment from the HSV ICP4 promoter and subjected to electrophoresis on a non-denaturing 5% polyacrylamide gel in .5X TBE (Garner et al, Nucl. Acids Res. 9, 3047 (1981); Fried et al, *ibid*, p. 6505 (1981)). For DNA binding assays, samples containing <u>in vitro</u> translated protein were incubated in 25 mM Tris pH 8.0, 10% glycerol, 50 mM CKl, 3 mM MgCl2, 0.5mM EDTA, 1mM DTT, 50 µg/ml poly dIdC on ice for 10 minutes, then probe was added and incubation continued at room temperature for 10 minutes. The probe was a 180 bp Nco I-Sal I fragment excised from the herpes simplex virus ICP4 promoter. The fragment was labeled by fill-in with the Klenow fragment of DNA polymerase I in the presence of $^{32}$P-dCTP. Protein:DNA complexes were subjected to electrophoresis on 5% (30:1) polyacrylamide gels in 0.5X TBE. Radioactive DNA and DNA:protein complexes were visualized by autoradiography. "B" indicates GABP bound DNA, "E" indicates DNA bound by proteins endogenous to reticulocyte lysates.

Figure 5. Sch matic diagram of GABP subunits showing regions of amino acid sequence similarity to

related proteins. (Top) GABPα is represented as a rectangle with the $NH_2$-terminus on the left and the COOH-terminus on the right. The region of sequenc similarity to Ets-related proteins is shaded (amino acids 316-400) and compared with the sequences of Ets-1 (Gunther et al, Genes Dev. 4, 667 (1990)), Erg (Reddy et al, Proc. Natl. Acad. Sci. U.S.A. 84, 6131 (1987)), Elk (Rao et al, Science 244, 66 (1989) and E74A (Burtis et al, Cell 61, 85 (1990). Residues that are common to GABPα and other proteins are boxed in black. (Bottom) GABPβ1 is represented as a rectangle with the $NH_2$-terminus on the left and the COOH-terminus on the right. The 33-amino acids repeats are shown as shaded rectangles. The unique COOH-terminal segment of GABPβ1 relative to GABPβ2 is indicated in black (333-382). The sequence of the four 33 amino acid repeats in GABPβ1 are shown below; residues that are common to two or more repeats are boxed in black and used to derive the GABPβ consensus. Similar criteria were used to derive consensus sequences for the 33 amino acid repeats of cdc 10/SW14,6 (Ares et al, EMBO J. 4, 457 (1985); Andrews et al, Nature, 342, 830 (1989); Breeden et al, Nature 329, 651 (1987)), Notch (Wharton et al, Cell 43, 567 (1985); Greenwald, *ibid*. 583 (1985)), glp1 (Yochem et al, Cell 58, 553 (1989); Austin et al *ibid*. p. 565 (1989)), lin12 (Yochem et al, Nature 335, 547 (1988)), ankyrin (Lux et al *ibid*. 344, 36 (1990)), NF B Kieran et al, Cell 62, 1007 (1990); Bours et al, Nature 348, 76 (1990)), fem1 (Yochem et al, Cell 58, 553 (1989; Austin et al *ibid*., p. 565 (1989)), and bcl-3 (Ohno et al *ibid*., p. 991 (1990)). The repeats from cdc10, SW14 and SW16 were combined to determine the consensus. The consensus for ankyrin was taken from Lux et al, Nature 344, 36 (1990). The overall consensus was defined as residues present in at

least 6 of the individual consensus repeat sequences.

Figure 6. DNA binding by GABP expressed in bacteria. Purified proteins were incubated with a $^{32}$P-labeled oligonucleotide containing the GABP binding site derived from the enhancer of the herpes simplex virus ICP4 gene (5'-TGCGGAACGGAAGCGGAAACCGCCGGATCG-3') (Triezenberg et al, Genes Dev. 2 (1988); LaMarco et al, *ibid.* 3, 1372 (1989)). Free and protein-bound DNA samples were subjected to electrophoresis on 5% polyacrylamide gels in either 0.5X TBE (A) or 0.25X TBE buffer (B).

Figure 7. Characterization of the DNA binding site for GABP. (A) Increasing concentrations of GABPα, either in the absence (left panel) or presence (right panel) of GABPβ1 were mixed with a $^{32}$P-labeled DNA fragment derived from the herpes simplex virus ICP4 enhancer. Free and protein-bound complexes were partially digested with DNase I and subjected to electrophoresis on an 8% polyacrylamide sequencing gel. The positions of three purine-rich repeats within the region of DNA protected from digestion by GABP are indicated by arrows. Lanes 1-6 (left panel) show digestion patterns resulting from GABPα concentrations starting at 1.5 nM and decreasing in 3-fold increments to 0.005 nM. Lanes 1-6 (right panel) show patterns resulting from addition of the same concentrations of GABPα that had been supplemented with 0.5 nM of GABPβ1. (B) Methylation protection (left panel) and interference (right panel) assays of DNA binding by GABP. The same DNA fragment used in (A) was incubated with GABPα, GABPβ1, or an equimolar mixture of the two subunits, and exposed to dimethyl sulfate (DMS).

Partially m thylated DNA was recovered, cleaved with piperidin , and run on an 8% polyacrylamide sequencing gel. For methylation interference assays, DNA was partially methylated, incubated with an equimolar mixture of GABPα and GABPβ1, and subjected to electrophoresis on a non-denaturing polyacrylamide gel as described in Fig. 6. Free and protein bound DNA species were recovered, cleaved with piperidine, and electrophoresed on an 8% polyacrylamide sequencing gel. Nucleotide residues closely contacted by GABP are shown in the lower part of (B). Filled circles identify guanine residues that were protected from DMS by GABP. Methylation of the same four guanine residues also inhibited DNA binding by GABP. Open circles identify adenine residues where methylation is enhanced in the presence of both GABPα and GABPβ1.

Figure 8. Measurements of DNA binding stability of complexes formed by various mixtures of GABP subunits. A $^{32}$P-labeled oligonucleotide containing a GABP binding site (Fig. 6) was incubated with GABPα alone, or together with equimolar amounts of either of the two β subunits. After a 10 minute incubation at 24°C, protein:DNA complexes were challenged with a 500-fold excess of unlabeled oligonucleotide. Protein-bound and free DNA were separated by non-denaturing gel electrophoresis as described in Fig. 6A. Protein-bound and free DNAs were located by autoradiography, excised, and quantitated by scintillation spectrometry. Results are presented as fraction of probe bound, normalized to 1.0 at the start point (time = 0).

Figure 9. UV-mediated crosslinking of GABP subunits to DNA. Isolated or mixed GABP subunits were incubated with a $^{32}$P-labeled oligonucleotide

containing a GABP binding site (Chodosh in Current Protocols in Molecular Biology, Vol. II, Ausubel et al eds (Greene Wiley, New York, 1988)) then exposed to ultraviolet light for varying lengths of time. UV crosslinking was performed using an oligonucleotide composed of a GA binding site flanked by 10 bp of non-specific sequence (5' AACCAAGCTTGCGGAACGGAAGCGGAAACCG 3') corresponding to residues located between 280 and 300 bp upstream of the herpes simplex virus gene encoding ICP4. Oligonucleotides were labeled to high specific activity by fill-in reaction with the Klenow fragment of DNA polymerase I in the presence of all four $^{32}$P-labeled dNTPs. DNA binding reactions were performed as described in a 96-well culture dish, followed by exposure to ultraviolet light. Samples were boiled in SDS-sample buffer and subjected to electrophoresis on SDS-polyacrylamide gels. Crosslinked protein species were visualized by autoradiography. Samples were denatured by boiling in SDS sample buffer and subjected to electrophoresis on a denaturing 12.5% polyacrylamide gel. Following electrophoresis the gel was dried and exposed to X-ray film. Time of exposure to UV light (minutes) is indicated above each gel lane. Migration of molecular weight marks (kD) is shown on the left.

Figure 10. Glutaraldehyde crosslinking of GABPß1 and GABPß2 subunits. Bacterially synthesized proteins were incubated in phosphate buffered saline with varying concentrations of glutaraldehyde as indicated below each lane for five minutes at room t mperature. Samples were denatured by boiling in SDS sample buffer and subject to electrophor sis on a denaturing 10% polyacrylamid gel. Following electrophoresis the gel was stained with Coomassie

brilliant blue. Proteins present in crosslinking reactions are indicated above each lane. ßN110 is a truncated version of GABPß1 missing 110 $NH_2$-terminal residues (see Fig. 12B).

Figure 11. DNA binding and complex formation assays of deleted variants of GABPα. Top panel shows schematic representation of GABPα deletion mutants. Individual mutants are designated according to the position of deletion end points with respect to the amino acid sequence of GABPα. Prefix "N" designates deletions missing residues starting at the $NH_2$-terminus of GABPα, prefix "C" designates deletions missing COOH-terminal residues, numbers indicate the position of the amino acid at which the deletion terminates. The Ets-related segment of GABPα is highlighted by grey stippling. Bottom panel shows an autoradiographic image of a non-denaturing gel used to separate DNA:protein complexes formed between variants of GABPα, GABPß1 and a $^{32}P$-labeled oligonucleotide that contained a GABP binding site. Each variant of GABPα was tested for DNA binding in the absence and presence of GABPß1 as indicated above the individual lanes.

Figure 12. Complex formation and UV crosslinking assays of deleted variants of GABPß1. Top panels of (A) and (B) show schematic representations of GABPß1 deletion mutants. Individual mutants are designated according to the positions of deletion end points with respect to the amino acid sequence of GABPß1. Prefix "N" designates deletions from the $NH_2$-terminus of GABPß1 (B), prefix "C" designates deletions missing COOH-terminal residues. Repeated sequences 33 or 32 amino acids in length that are related to similarly sized repeats in th Notch protein of Drosophila melanogaster are highlighted

by grey stippling. Unique parts of GABPβ1 and GABPβ2 are indicated by black and hatched rectangles at their respective COOH-termini. Deleted formation with GABPα as shown in the lower left panels of (A) and (B). Each deletion mutant was also tested in UV crosslinking assays shown in the lower right panels of (A) and (B). All complex formation and UV crosslinking assays were conducted in the presence of GABPα and a $^{32}$P-labeled oligonucleotide containing a GABP binding site.

Figure 13. Model depicting complex formed between GABP and DNA. The sequence of the GABP binding site consists of two hexanucleotide repeats of the sequence 5'-CGGAAR-3' as in lower part of Fig. 13. Oval spheres directly above guanine residues of each hexanucleotide correspond to GABPα subunits, elongated rectangles correspond to 33 amino acid repeats of GABPβ subunits. Smaller rectangles shown at top correspond to the region of GABPβ1 required for formation of stable homodimers. Circular arrows designate flexible regions inferred to occur between the dimer forming region of GABPβ1 and the 33-amino acid repeats located at its $NH_2$-terminus.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to a DNA segment encoding all (or a unique portion) of the heteromeric transcriptional regulatory protein termed GA binding protein (GABP). The invention further relates to the encoded proteins (or polypeptides). A "unique portion" as used herein consists of at least five (or six) amino acids or, correspondingly, at least 15 (or 18) nucleotides. The present invention further relates to a

recombinant DNA mol cule comprising the above DNA segment and to host cells transformed therewith.

In particular, the present invention relates to a DNA segment that encodes the entire amino acid sequence of GABPα, GABPβ1 or GABPβ2 given in Figure 2 (the specific DNA segments given in Figure 2 being only examples), or any unique portion thereof. DNA segments to which the invention relates also include those encoding substantially the same protein subunits as shown in Figure 2, including, for example, allelic and species variations thereof and functional equivalents of the amino acid sequences of Figure 2. The invention further relates to a DNA segment substantially identical to one of the subunit sequences shown in Figure 2. A "substantially identical" sequence is one the complement of which hybridizes to one of the sequences of Figure 2 at 50°C and 6X SSC (saline/sodium citrate) and which remains bound when subjected to washing at 55°C with 6X SSC (note: 20 x SSC = 3M sodium chloride/0.3 M sodium citrate). The invention also relates to nucleotide fragments complementary to such DNA segments. Unique portions of the DNA segment, or complementary fragments, can be used as probes for detecting the presence of respective complementary strands in DNA (or RNA) containing samples.

The present invention further relates to GABP, and subunits thereof, substantially free of proteins with which it is normally associated, and more especially, to unique peptide fragments of the subunits of that protein. The GABP protein (or functionally equivalent variations thereof), or p ptide fragments thereof, to which the invention relates, also includes those which are chemically synthesized using known methods. The proteins and

peptides of the present invention can be modified, for example, phosphorylated, or unmodified.

The present invention also relates to recombinantly produced GABP, or subunits thereof, having the amino acid sequence shown in Figure 2 or functionally equivalent variation thereof. The recombinantly produced protein can be modified, for example phosphorylated, or unmodified. The present invention, more particularly, relates to recombinantly produced unique peptide fragments of GABP subunits.

The present invention also relates to a recombinant DNA molecule (or construct) and to a host cell transformed therewith. Using standard methodologies, well known in the art, a recombinant DNA molecule comprising a vector and a DNA segment encoding at least one GABP subunit, or a unique portion thereof, can be constructed. Vectors suitable for use in the present invention include plasmid and viral vectors. The vector can be selected so as to be suitable for transforming prokaryotic or eukaryotic cells. Advantageously, the recombinant molecule includes a promoter operably linked to the GABP encoding segment.

The recombinant DNA molecule of the invention can be introduced into appropriate host cells by one skilled in the art using method well known in the art. Suitable host cells include prokaryotic cells, such as bacteria, lower eukaryotic cells, such as yeast, and higher eukaryotic cells, such as mammalian cells. These cells can serve as a source of GABP when cultured under appropriate conditions.

As noted at the outset, and as will be further described in the Examples that follow, significant amino acid sequence similarity exists between GABP$\alpha$ and the products of the *est1* and *ets2*

(Watson et al, Proc. Natl. Acad. Sci. U.S.A. 85, 7862 (1988); Gunther et al, Genes Dev. 4, 667 (1990)) proto-oncogenes. The Ets-related r gion of GABPα is located close to the COOH-terminus of the subunit. Biochemical studies of Ets1 (Gunther et al, Genes Dev. 4, 667 (1990)), as well as the related proteins PU (Klemsz et al, Cell 61, 113 (1990)) and E74 (Urness et al, Cell 63, 47 (1990)), have demonstrated direct, sequence-specific DNA binding. These proteins, as well the products of several additional eukaryotic genes, share sequence similarity in an 85-amino acid region that is required for DNA binding (Karim et al, Genes Dev. 4, 1451 (1990)). The region of GABPα that is related to this family of proteins coincides with the 85 amino acid DNA binding domain (Fig. 5).

The amino acid sequences of GABPβ1 and GABPβ2 contain four repeats of a related amino acid sequence located at the $NH_2$-termini of both subunits (Fig. 5). The first two repeats are 32 amino acids in length and the second two contain 33 amino acids. Similar repeats occur in the Notch protein of Drosophila melanogaster (Wharton et al, Cell 43, 567 (1985); I. Greenwald, *ibid.*, 583 (1985)), and the Lin12 and Glp1 proteins of Caenorhabditis elegans (Yochem et al, Nature 335, 547 (1988), Yochem et al, Cell 58, 553 (1989); Austin et al, *ibid.*, p. 565 (1989)). These "33-amino acid repeats" were first recognized in studies of the yeast protein SW16, which regulates gene expression involved in mating type switching (Breeden et al, Nature 329, 651 (1987)). Similar repeats have been identified in ankrin, a multifunctional protein associated with the membrane of red blood cells (Lux et al, Nature 344, 36 (1990)), several vaccinia virus encoded proteins of unknown function (Gillard et al, Proc. Natl. Acad. Sci. U.S.A. 83, 5573 (1986)), and the

transcription factor NFKB (Kieran et al, Cell 62, 1007 (1990); Bours et al, Nature 348, 76 (1990)).

The two subunits of GABP exhibit primary sequence motifs typical of proteins normally found in different cellular compartments. Accordingly, transcriptional regulatory proteins, such as members of the Ets family, might interact with membrane bound proteins that contain the 33-amino acid repeats present in GABPß. The Notch, Glp1 and Lin12 proteins might sequester transcription factors at the plasma membrane which could be released in response to appropriate extracellular signaling events. Alternatively, the cytoplasmic segments of these transmembrane proteins might be proteolyzed in response to an extracellular signal, allowing the 33-amino acid repeats to be translocated to the nucleus where they could abet the action of a second subunit. Either scenario would offer a direct pathway of signal transduction.

The results set forth in the Examples that follow demonstrate the reliance of competent DNA binding complexes on multiple subunits. By separating functional components onto different polypeptide chains, critical subunits might be differentially expressed or sequestered, generating useful strategies for regulation. For example differential expression of the mRNAs encoding the ß1 and ß2 subunits of GABP would be expected to impact substantially on the function of the resulting complex. In this regard, it is interesting to note that in cells undergoing replication, the ß2 subunit predominates whereas in non-dividing cells, subunit ß1 predominates.

It will be clear to one skilled in the art from a reading of this disclosure that advantage can be taken of information provided herein to effect alterations of both normal and abnormal expression

patterns regulated by the binding complexes described above. Such alterations can be effected, for example, using a vari ty of gen therapy protocols. It is contemplated that by altering the relative amounts of the ß1 and ß2 subunits, disease states characterized by rapid cell division, for example, cancer, can be controlled.

Certain aspects of the invention are described in greater detail in the non-limiting Examples that follow.

## Example 1

### Isolation of Recombinant cDNA Clones

GABP (20 µg) was purified from rat liver nuclear extracts and cleaved with trypsin. Proteolyzed fragments were separated by high performance liquid chromatography (HPLC), recovered, and subjected to gas-phase amino acid sequencing (Fig. 1). Partial sequences were derived from 13 tryptic peptides. Degenerate oligonucleotides capable of encoding four of the thirteen peptide sequences were synthesized and used as hybridization probes to screen an adipocyte cDNA library. Degenerate oligonucleotides were labeled with $^{32}P$ using polynucleotide kinase. The basic sodium chloride/sodium citrate (SSC) protocol was used for screening (F.M. Ausubel et al, Current Protocols in Molecular Biology (Wiley and Sons, NY), 1989). Hybridization was performed at 48°C for 16 hours. One recombinant bacteriophage that contained a cDNA insert of 2 kb gave a positive signal when hybridized with each of the four oligonucleotide probes.

The insert of this r combinant was sequenced and found to contain an opening reading

frame that encoded a protein of 454 amino acids (Fig. 2A). The predicted molecular weight of this polypeptide (51.3 kD) corresponded to the size of the GABPα subunit purified from rat liver nuclei (LaMarco et al, Genes Dev. 3, 1372 (1989); Fig. 1). Inspection of the deduced amino acid sequence revealed segments that corresponded to eight of the 13 peptides isolated by trypsin digestion of intact GABP. On the basis of the latter two observations, this 454 residue polypeptide was tentatively identified as GABPα.

Degenerate oligonucleotides capable of encoding two of the tryptic peptide sequences not present in GABPα were synthesized and used as hybridization probes to search for a cDNA clone that encoded GABPβ (Lee, Mol. Endocrinol. 4, 1034 (1990)). Kinased oligonucleotide probes were hybrided in 6X SSC, 1X Denhardt's, 0.05% sodium pyrophosphate, and 100 μg/ml yeast tRNA at 50°C for 14 hours. Washing conditions were 6X SSC, 0.05% sodium pyrophosphate at 55°C. A total of five clones were isolated that hybridized with both oligonucleotide probes. Four of the clones were approximately 2.6 kb and differed only slightly in the length of the 5' untranslated region; these cDNA clones encoded GABPβ1. The fifth cDNA clone was approximately 1.4 kb and differed from the other four at its 3' end; this cDNA clone encoded GABPβ2. Four additional cDNA clones corresponding to GABPβ2 were subsequently identified. Five recombinant bacteriophage were identified according to their capacity to hybridize with both oligonucleotide probes. One of the cDNA clones differed at the 3' end from th other four. The largest cDNA insert of the four (2.6 kb) and the variant (1.4 kb) were sequenced (Fig. 2B). Both DNA sequences r vealed long open reading frames specifying highly similar

polypeptid s. One cDNA encoded a protein of 382 amino acids, the other encoded a protein of 349 residues. Starting at their respective $NH_2$-termini, the two proteins exhibited identical sequences for 333 amino acids. At this point the sequences to the two protein diverged such that the longer one contained an additional 50 residues before its terminus. The divergent COOH-terminal segments bore no apparent amino acid sequence similarity. The open reading frames of both polypeptides contained segments that corresponded to the two tryptic peptides used to design hybridization probes. Moreover, the predicted molecular weights of the two polypeptides (41.3 and 37 kD) corresponded closely with the size of the GABPβ subunit purified from rat liver nuclei (LaMarco et al, Genes Dev. 3, 1372 (1989); Fig. 1). The 41 kD polypeptide was therefore provisionally designated as GABPβ1 and the 37 kD polypeptide as GABβ2.

## Example 2

### Tissue Distribution of mRNA Encoding GABPα, GABPβ1 and GABPβ2

Northern (RNA) blot assays were used to determine the sizes and tissue distributions of mRNA encoding GABPα, GABPβ1 and GABPβ2 (Fig. 3). The cDNA corresponding to GABPα identified three mRNAs of roughly 5.0, 2.8 and 2.6 kb, which were expressed in a variety of tissues. The GABPα cDNA, which consisted of slightly less than 2.0 kb (Fig. 2A), represents a partial copy of any of the three mRNAs. Two mRNAs measuring 2.7 and 1.5 kb were identified in northern blots probed with GABPβ1 cDNA. Like GABPα mRNAs, those encoding GABPβ1 had a wide tissu distribution. Because the cDNAs that encoded GABPβ1

and GABPβ2 measured 2.6 and 1.4 kb, respectively (Fig. 2B), they probably represent nearly full-length copies of the respective mRNAs. Moreover, because the nucleotide sequences of the two cDNAs are identical from their respective 5' termini to the point of abrupt divergence 1.1 kb internal to the mRNA, they likely represent alternatively spliced transcripts derived from the same gene. Consistent with this interpretation is the presence of a potential splice donor site (AG dinucleotide) immediately preceding the point of divergence.

## Example 3

### GABP DNA Binding Activity

To test whether the recombinant DNA clones described above possessed GABP DNA binding activity, reticulocyte lysates were programmed with RNA synthesized from the cDNAs that encode GABPβ, GABPβ1 and GABPβ2. Each RNA was translated to form a protein product of the expected size (Fig. 4A). Individual lysates or mixtures thereof were tested for DNA binding to a fragment from the HSV1 ICP4 promoter that contained three GA repeats. Protein:DNA mixes were subjected to electrophoresis on nondenaturing polyacrylamide gels to separate free DNA from that complexed with protein. Reticulocyte lysate that had not been programmed with exogenous RNA contained protein(s) capable of forming a complex with the oligonucleotide probe that migrated more rapidly than the complex formed by GABP. Other than background activity endogenous to the reticulocyte lysate, specific protein:DNA complexes were not observed when lysates programmed with GABPα, GABPβ1, or GABPβ2 were tested in electrophoretic mobility shift assays. Likewise, no

new DNA binding activity was observed with lysat that had been us d to co-translate RNAs encoding GABPα and GABPβ2. However, co-translation of RNAs encoding GABPα and GABPβ1 caused the lysate to form a DNA binding activity that could be distinguished from background (Fig. 4B). The interdependency of GABPα and GABPβ1 observed in these assays is consistent with earlier observations that tested subunits purified from rat liver nuclei (LaMarco et al, Genes Dev. 3, 1372 (1989)).

## Example 4

### DNA Binding Properties of GABP

Recombinant cDNA copies of the mRNAs that encode GABPα and GABPβ1 were introduced into bacteriophage T7 based vectors that allowed synthesis of the corresponding proteins in Escherichia coli (Studier et al, J. Mol. Biol., 189, 113 (1986)). Polymerase chain reaction was used to introduce a Bam HI site at the 5' end of the open reading frames encoding GABPα or GABPβ1. cDNAs lacking the 3' untranslated region were inserted into a modified pT5 vector, which adds two amino acids (gly-ser) at the $NH_2$-terminus of the encoded protein. Each subunit was expressed and purified using conventional chromatographic techniques. Proteins were expressed in bacteria as described (Shuman et al, Science 249, 771 (1990)). GABPα was precipitated from the soluble fraction by the addition of one volume of 2M ammonium sulfate in buffer A (10 mM Tris-HCl, pH 7.9, 100 mM KCl, 0.5 mM EDTA, 1 mM $MgCl_2$, 2 mM DTT, 0.1 mM phenylmethylsulfonyl fluoride) with 2 mM $CaCl_2$. The ammonium sulfate pellet was resuspended in 25 ml buffer B (25 mM Tris-HCl, pH 8.0, 0.75 mM EDTA, 10%

(v/v) glycerol, 1 mM DTT) with 75 mM NaCl and dialyzed against the same buffer. The dialysate was loaded onto a column of Q-Sepharose Fast Flow (Pharmacia). GABPα was eluted with a 75-500 mM NaCl gradient in buffer B. Peak fractions were pooled, dialyzed against buffer B and loaded onto a salmon sperm DNA-sepharose column. GABPα was eluted with a 0-400 mM NaCl gradient. GABPα was judged by Coomassie Blue staining of SDS polyacrylamide gels to account for greater than 90% of the total protein. GABPβ1 was solubilized from the particulate fraction of bacterial extracts by sonication in buffer A supplemented with 7 M urea. The urea solubilized fraction was dialyzed against buffer B with 75 mM NaCl and centrifuged at 16,300 x g for one hour. The supernatant was applied to a Q-Sepharose column and eluted with a gradient of 75-500 mM NaCl. GABPβ1 was judged to account for greater than 90% of total protein by Coomassie Blue staining of SDS-polyacrylamide gels.

The DNA binding properties of the two individual polypeptides and mixtures thereof were first studied by gel retardation using a DNA substrate derived from the enhancer of an immediate early gene of herpes simples virus. Consistent with earlier studies (LaMarco et al, Genes Dev., 3, 1372 (1989)), binding was not observed when DNA was incubated with either of the isolated subunits. When GABPα and GABPβ1 were incubated with DNA simultaneously, a DNA:protein complex exhibiting substantially retarded mobility relative to free DNA was observed (Fig. 6, left panel).

The multi-subunit dependence of DNA binding by GABP was relieved in gel retardation assays conducted at lower ionic strength (Fig. 6, right panel). GABPα formed two retarded complexes that migrated at positions between free DNA and the

complex formed with both subunits. Under these conditions, th mixture of GABPα and GABPβ1 again led to the formation of a slowly migrating DNA:protein complex. The retarded mobility of the latter complex, relative to those formed by GABPα alone, reflected the presence of the GABPβ1 subunit. The validity of this interpretation was confirmed by the use of antisera specific to each subunit. Antiserum specific to GABPα further retarded the migration of complexes formed with GABPα alone or the mixture of GABPα and GABPβ1. Antiserum specific to GABPβ1 did not affect the mobility of complexes formed between GABPα DNA, but retarded the complex formed in the presence of both subunits. Polyclonal antisera were generated by rejecting rabbits with purified GABPα or GABPβ1. Antisera were added to gel shift reactions at a dilution of 1:20. Pre-immune sera did not effect the migration of protein:DNA complexes.

The HSV1-derived DNA fragment used in binding assays of GABP contains three imperfect repeats of the hexanucleotide sequence 5'-CGGAAR-3', which were shown in earlier studies to be protected from DNase I digestion when bound by GABP (Triezenberg et al, Genes Dev. 2 (1988); LaMarco et al, *ibid*. 3 1372 (1989)). DNase I footprinting assays were performed using bacterially synthesized proteins under conditions that allowed interaction of GABPα alone. As shown in Fig. 7A, GABPα was capable of protecting the repeated hexanucleotide motifs from DNase I digestion when added at a concentration of 0.15 nM. When the GABPβ1 subunit was added, protection was observed at a 10-fold more dilute concentration of GABPα (0.015 nM). In addition, th pattern of nuclease protection was extended slightly beyond the adenine residues of the third r peat.

The segment of DNA protected from DNase I digestion by GABP encompassed all three hexanucleotide repeats, yet was not centered over the repeats. Methylation protection and interference assays were undertaken in order to gain a more refined image of the sites of close contact established between GABP and DNA. Methylation protection assays conducted with GABPα showed a pattern of protection that included both guanine residues of the second and third hexanucleotide repeats. The same sets of guanines were protected when GABPβ1 was added to the binding reaction. In the latter case, however, accentuated methylation was observed at adenine residues located adjacent to the guanine dinucleotides of the second and third repeats. Sites of methylation interference were mapped by separating protein-bound DNA molecules from those inactivated by partial methylation. Methylation of guanine dinucleotides in the second and third hexanucleotide repeats inhibited binding by the mixture of GABPα and GABPβ1 (Fig. 7B).

The results of methylation protection and interference assays indicate that GABP binds to sites on DNA corresponding to two of the three purine-rich hexanucleotide repeats. The pattern of protection of guanine residues by GABPα was similar to that observed with the mixture of both subunits, indicating that GABPα, when added at a sufficiently high concentration, can bind specifically to DNA in the absence of GABPβ1. Such observations offer an explanation for the two retarded bands observed when DNA was challenged with GABPα alone under conditions of low ionic strength (Fig. 6, right panel). The less retarded of the two bands is interpreted to represent a complex wherein GABPα is associated with only one of the two hexanucleotide repeats, while the more retarded complex is interpreted to contain

GABPα subunits associated with two hexanucleotide repeats. Binding assays that test d DNA probes containing a single hexanucleotide repeat supported this interpretation. When incubated with GABPα and assayed in low ionic strength gels, such DNA probes generated only one retarded complex.

Several observations indicate that the mixture of GABPα and GABPß1 forms a complex that binds DNA more stably than the α subunit alone (LaMarco et al, Genes Dev. 3, 1372 (1989)). To further investigate the effect of GABPß1 on DNA binding by GABPα, the rate at which variously mixed proteins dissociate from DNA was measured (Fig. 8). The dissociation rate of GABPα alone was too rapid to be accurately measured. Less than ten percent of the DNA remained bound to GABPα after a 10 second challenge with excess, unlabeled competitor DNA. In contrast, when both GABPα and GABPß1 were present, the dissociation rate was much slower ($T_{1/2}$ = 1.5 min). Similar assays were performed with a mixture of GABPα and GABPß2, which, in earlier experiments, failed to form a stable complex with DNA. When used at nM concentrations, GABPß2 was capable of forming a DNA binding complex with GABPα (see Fig. 12A). The ß2 isoform of GABP also stabilized DNA binding by GABPα, yet yielded a complex that dissociated more rapidly ($T_{1/2}$ = 30 s) than that formed with GABPß1. Because the ß1 and ß2 isoforms differ only at their COOH-termini, this part of the protein may be involved in stabilizing DNA binding.

The observations outlined thus far indicate that the ß1 and ß2 isoforms of GABP do not bind to DNA alone, but associate with the α subunit to augment DNA binding. The question arises, therefore, as to whether the ß subunits cause a conformational change in α leading to its more avid interaction with DNA. Alternatively, or in addition

to causing a conformational change, the ß subunits might, in association with GABPα, establish direct contact with DNA.

To determine whether GABPß1 contacted DNA when complexed with GABPα, DNA:protein complexes were exposed to ultraviolet (UV) light under conditions expected to permit covalent crosslinking between DNA and intimately bound proteins (L. A. Chodosh in Current Protocols in Molecular Biology, vol II, F. M. Ausubel et al., eds. (Greene/Wiley, New York, 1988). UV crosslinking was performed using an oligonucleotide composed of a GA binding site flanked by 10 bp of non-specific sequence ((5' AACCAAGCTTGCGGAACGGAAGCGGAAACCG 3') corresponding to residues located between 280 and 300 bp upstream of the herpes simplex virus gene encoding ICP4. Oligonucleotides were labeled to high specific activity by fill-in reaction with the Klenow fragment of DNA polymerase I in the presence of all four $^{32}$P-labeled dNTPs. DNA binding reactions were performed as described in a 96-well culture dish, followed by exposure to ultraviolet light. Samples were boiled in SDS-sample buffer and subjected to electrophoresis on SDS-polyacrylamide gels. Crosslinked protein species were visualized by autoradiography. The GABP subunits were incubated with $^{32}$P-labeled DNA that contained the purine-rich hexanucleotide repeats, exposed to UV light, and subject to electrophoresis on a denaturing polyacrylamide gel. When DNA was incubated with GABPα and exposed to UV light, a crosslinked product was observed bearing an electrophoretic mobility close to that of GABPα (Fig. 9). The appearance of this product was dependent on the presence of GABPα and increased in a time-dependent manner upon exposure to UV light. Moreover, it was eliminated by the inclusion of excess, unlabeled DNA that

contained the purine-rich hexanucleotide repeats, but not by excess non-specific DNA.

No evidence of protein:DNA crosslinking was observ d when $^{32}$P-labeled DNA was mixed with GABPß1 and exposed to UV light. However, when the mixture of GABPα and GABPß1 was complexed with DNA and irradiated, new crosslinked products were observed. In addition to GABPα, two closely migrating polypeptide bands, slightly larger than the native size of GABPß1, became covalently attached to the radioactive DNA substrate (Fig. 9). Although GABPß1 was incapable of binding DNA on its own, when present in a ternary complex it appeared to be even more susceptible to UV-mediated crosslinking than GABPα. These data provide evidence that the ß1 subunit of GABP associates closely with DNA when complexed with GABPα.

## Example 5

### Formation of a Stable Complex between GABPα and GABPß in the Absence of DNA

Having found that the α and ß subunits of GABP formed a heteromeric complex when exposed to their specific DNA substrate, gel filtration chromatography was used to determine whether these subunits might associate in the absence of DNA. Gel filtration chromatography was performed using a Superose-6 column (10 x 300 cm, Pharmacia) in buffer B supplemented with 0.4M NaCl. The column was calibrated with molecular weight standards thyroglobulin, apoferritin, catalase, bovine serum albumin and ribonuclease. 50-100 μg of each protein was chromatographed at 0.5 ml/min. Elution volume was converted to $K_{av}$ by the equation $K_{av} = (V_e-V_o)/V_t-V_o)$ where $V_o$ = void volume = 8.1 ml; $V_t$ + total bed

volume = 24.0 ml; V = eluted volume. The Stokes radius was calculated from a plot of $(-\log K_{av})^{1/2}$ versus Stokes radius (Ackers, Adv. Prot. Chem. 24, 343 (1970)). GABPα eluted as a single peak at 15.2 ml; GABPβ1 at 14.0 ml; GABPβ2 at 15.8 ml. A mixture of equal amounts of GABPα and β1 chromatographed as a single peak at 12.1 ml. The mixture of GABPα and GABPβ2 chromatographed as a single peak at 13.9 ml. Loaded separately, GABPα and GABPβ1 eluted as single peaks at $K_{av}$ values of 0.45 and 0.38 respectively. However, when loaded together at an equimolar ratio, both subunits eluted as a single peak at $K_{av}$ 0.26 (Table 1). Analysis of column fractions by polyacrylamide gel electrophoresis confirmed that the peak at $K_{av}$ 0.26 contained both subunits.

TABLE 1

Determination of molecular weights of GABP subunits. Purified GABPα, GABPβ1 and GABPβ2 produced in E. coli were analyzed by gel filtration and sedimentation velocity as described (17,18). $K_{av}$ values were calculated from the elution volume of a Superose-6 FPLC column. Apparent molecular weights were determined from $K_{av}$ vs. log MW for the column. Stokes radii were determined from a plot of $(-\log K_{av})^{\frac{1}{2}}$. Sedimentation coefficients together with measured Stokes radii were used to calculate native molecular weights.

| Protein | $K_{av}$ | Apparent MW ($K_{av}$ vs. log MW) | Stokes Radius | Sedimentation Coefficient | Corrected MW |
|---|---|---|---|---|---|
| GABPα | 0.449 | 158,000 | 49.3 | 3.1 | 66,000 |
| GABPβ1 | 0.377 | 349,000 | 63.0 | 3.1 | 82,000 |
| GABPβ2 | 0.486 | 106,000 | 44.0 | 2.5 | 46,000 |
| GABPα + GABPβ1 | 0.255 | $1.3 \times 10^6$ | 87.7 | 4.5 | 170,000 |
| GABPα + GABPβ2 | 0.367 | 390,000 | 64.8 | 3.8 | 104,000 |

Evidence confirming the stable association of GABP subunits in the absence of DNA was also obtained from measurements of sedimentation velocity (Table 1). The gel filtration and sedimentation properties of a protein or protein complex are affected both by size and molecular shape. However, by using the analytical methods of Siegel and Monty (Martin et al., J. Biol. Chem. 236, 1372 (1961)), it was possible to calculate native molecular weights of the various protein species. Sedimentation coefficients were determined on 4.5 ml 10-30% glycerol gradients in 25 mM Tris-HCl pH 8.0, 75 mM NaCl, 0.75 mM EDTA, 1 mM DTT. 30 μg of each protein was loaded in 0.1 ml together with catalase, bovine serum albumin, and cytochrome c as internal standards. Gradients were centrifuged at 4°C for 40 hours at 39,000 rpm. Fractions (0.25 ml) were collected and analyzed by SDS-PAGE with Coomassie blue staining. The S value for each sample was determined by its sedimentation relative to the BSA and cytochrome c standards. Native molecular weights were derived using the Stokes radius together with measured sedimentation coefficients as described in Siegel et al, (Biochim. Biophys. Acta 112, 346 (1966)). Partial specific volume was calculated using the predicted amino acid sequences of each GABP subunit as described by Cohn et al, Proteins, Amino Acids and Peptides as Ions and Dipolar Ions, E. J. Cohn and J. T. Edsall, eds. (Reinhold Publishing Co., New York, 1943), pp. 370-381). As shown in Table 1, the calculated molecular weights of GABPα and GABPβ2 corresponded closely to their predicted sizes (51.3 and 37 kD, respectively). In contrast, GABPβ1 exhibited a native molecular weight (82 kD) roughly twice its expected size (41.3 kD).

The complex formed between GABPα and GABPβ1 eluted from the gel filtration column prior to the largest molecular weight standard. Thus, the calculated molecular weight of this complex (170 kD) represents a provisional assignment. Since GABPβ1 existed as a stable dimer on its own, the very large complex formed between it and GABPα was tentatively identified as a tetramer composed of two molecules of each subunit. Polyacrylamide gel analysis of the constituents of the GABPα:GABPβ1 complex were consistent with this interpretation, showing that the subunits existed in equal stoichometries. This interpretation was also consistent with the properties of the complex formed between GABPα and the β subunits was interpreted to reflect the dimer forming property of GABPβ1.

Gel filtration and gradient sedimentation assays indicated that GABPβ1 might exist as a dimer. This interpretation was tested using glutaraldehyde crosslinking assays. Bacterially expressed GABPβ1 and GABPβ2 were exposed to glutaraldehyde and subjected to electrophoresis on a denaturing polyacrylamide gel. Incubation of GABPβ1 with glutaraldehyde led to the formation of a second polypeptide band exhibiting an apparent molecular weight roughly double that of the monomeric form of the protein (Fig. 10). Similar experiments conducted with GABPβ2 failed to yield an analogous product.

Additional evidence suggesting that GABPβ1 exists as a dimer resulted from crosslinking experiments with the intact polypeptide and a truncated form lacking the 110 $NH_2$-terminal residues (βN110, see Fig. 12B). Crosslinking of the truncated protein led to the formation of an additional species roughly double the size of the monomeric form. When the truncated protein was

mixed with intact GABPβ1 and exposed to glutaraldehyde, three crosslinked protein species were observ d. Two species corresponded to crosslinked, homodimeric complexes that had been observed upon glutaraldehyde treatment of active GABPβ1 and the $NH_2$-terminal truncated derivative. The third species migrated between the presumed homodimeric forms and probably represented a heteromeric complex consisting of one GABPβ1 polypeptide and one truncated polypeptide. Therefore, both molecular weight measurements and crosslinking assays showed that GABPβ1 but not GABPβ2 exists as a stable homodimer.

Example 6

Mapping of Functional Domains of GABPα and GABPβ

Experimental results described above indicate that GABPα should contain at least two functional components, one that facilitates DNA binding and another that allows complex formation with GABPβ. The GABPβ1 polypeptide should contain at least three components, facilitating self-dimerization, heterodimerization with GABPα, and direct contact with some part of the purine-rich DNA substrate. Recombinant copies of the genes that encoded each subunit were systematically deleted to localize these components. Deletion mutants of GABPα were generated by polymerase chain reaction and expressed in pT5 as described (Breeden et al, Nature 329, 651 (1987)). Soluble bacterial extracts containing deleted variants of GABPα were used for binding reactions. $NH_2$-terminal deletions of GABPβ1 were generated by exonuclease III digestion, followed by digestion with S1 nuclease and ligation of Bam HI link rs. All deletions were sequenced and

subcloned into the appropriate pET3 vector (Rosenberg et al., Gene 56, 125 (1987)) to maintain the proper reading frame. COOH-terminal deletions were generated using 3' deletions of the cDNA inserted in Bluescript (Stratagene) by subcloning Bst EII-Asp 718 or Sac I-Asp 718 fragments into the pET-GABPβ1 plasmid that had been digested with the appropriate enzymes. Translation termination codons were provided by vector sequences so that in some cases extra amino acids are appended to the open reading frame. All GABPβ derivatives were insoluble and re-solubilized in 8M urea followed by dialysis against 10 mM Tris pH 8.0, 75 mM KCl or NaCl, 1 mM DTT, 0.2mM PMSF, 1mM benzamidine, 10% glycerol prior to use in binding reactions. All derivatives were expressed equivalently as determined by Coomassie staining of SDS polyacrylamide gels.

Deletion variants of GABPα that were missing as many as 313 residues from the $NH_2$-terminus retained the capacities to bind DNA and complex with GABPβ (Fig. 11). A GABPα variant further missing 17 residues (αN313/437) from the COOH-terminus also retained both functions. More extensive deletion from the COOH-terminus, to amino acid 407 (αN313/C407), yielded a protein that was capable of binding to DNA, but had lost the ability to complex with GABPβ1. These results showed that the Ets-related domain of GABPα was sufficient for DNA binding. The region of GABPα required to form a complex with GABPβ included the Ets-related segment, as well as 37 amino acids located on the immediate COOH-terminal side of the Ets-related domain.

In order to define regions of GABPβ1 that interact with GABPα and contact DNA, systematically deleted variants were produced and tested in gel retardation and UV-crosslinking assays (Fig. 12). Variants that lacked up to 228 residues (βC154) from

the COOH-terminus of GABPβ1 proved to be functional in both assays. Deletion of an additional 33 residues (βC121) yielded a protein that failed to function in either assay. The boundary defined by these experiments corresponded to the location of the most COOH-terminal of four 33-amino acid repeats that are present in both isoforms of GABPβ.

Although about 70% of GABPβ1 could be deleted from its COOH-terminus without eliminating interaction with GABPα and DNA, removal of only a small segment from the $NH_2$-terminus resulted in deleterious effects. A variant of GABPβ1 that lacked 19 $NH_2$-terminal residues (βN19) was slightly less effective in converting GABPα-derived complexes into the very slowly migrating heteromeric complex. When tested in the UV crosslinking assay, βN19 yielded a reduced amount of crosslinked product relative to the intact β1 isoform. Variants that lacked 47 and 67 residues (βN47 and βN67) were progressively more defective in the complex formation assay and failed to be crosslinked to the radioactive DNA probe as efficiently as the intact protein. Finally, variants missing 80 or more residues from the $NH_2$-terminus were completely defective in both assays. The progressive loss of function observed in deleted forms of GABPβ1 corresponded to the progressive loss of the 33-amino acid repeats (Fig. 12). βN19 was truncated within the first of the repeats, βN47 within the second, βN67 after the second, and βN80 within the third repeat. The functional properties of the GABPβ1 deletion mutants indicate that the 33-amino acid repeats are important both for complex formation with GABPα and DNA contact.

## Example 7

### A Model for DNA-Bound GABP

The foregoing observations have been incorporated into a provisional model of the complex formed when GABPα and GABPß1 associate with a directly repeated set of purine-rich hexanucleotides (Fig. 13). Each hexanucleotide repeat is hypothesized to be contacted by both GABPα and GABPß1. The linear order of contact, wherein GABPα is associated with guanines on one side of each hexanucleotide and GABPß with adenines on the other side, was deducted from three separate observations. First, GABPα was alone capable of protecting both guanines from methylation by dimethysulfate. Second, the DNase I footprint generated by the mixed subunits, relative to that resulting from GABPα alone, was extended in a direction toward the adenine residues of the hexanucleotide repeat. Third, addition of GABPß caused an enhanced pattern of methylation of adenine residues relative to the pattern generated in binding reactions that contained GABPα alone.

The most stable complexes formed between GABP and DNA were observed with the mixture of GABPα and GABPß1. The ß1 subunit, unlike ß2, was observed to exist as a stable homodimer. Moreover, when mixed with GABPα, the ß1 subunit generated a high molecular weight complex probably consisting of two polypeptides of each subunit. This heteromeric tetramer is believed to bind in a concerted manner to two purine-rich hexanucleotide repeats. If that is the case, a flexible region should exist between the dimerization domain of GABPß1 and the surfaces located near its $NH_2$-terminus that facilitate interaction with GABPα and DNA. Without such

flexibility a linked set of polypeptides would not likely be capable of binding simultaneously to a DNA substrate that is not rotationally symmetric.

## Example 8

### Isolation of cDNA Clones Encoding Human GABP Alpha and Beta Subunits

cDNA clones encoding human GABP alpha, beta3 and beta4 were isolated by screening a human fetal brain cDNA library. cDNAs for human GABP beta1 and beta2 were isolated from a HeLa cell cDNA library. The probes used to screen for human GABP alpha were the 865 bp Ava 1-Sst 1 and 678 bp Bam H1-Sst 1 fragments of the mouse GABP alpha cDNA. The probe used to isolate the human beta was an 850 base pair fragment from the 5' end of the mouse GABP beta2 cDNA.

Purified DNA fragments used as probes were radiolabeled with 32P by random priming reactions. Hybridization conditions were 6X SSC, 1X Denhardt's, 0.05% sodium pyrophosphate, 100 μg/ml yeast tRNA. The final wash buffer was 2 X SSC. The hybridization and washing temperature were 65° C.

cDNAs were confirmed as the human homologs of mouse GABP by determination of their nucleotide sequence.

Six different cDNAs have been deposited with the American Type Culture Collection (ATCC), 12301 Parklawn Drive, Rockville, MD 20852.

1A = human GABPβ, Eco R1 fragment common to all beta isoforms, in Bluescript KS+, isolated from HeLa cDNA library.

4A = human GABPβ1 Eco R1 fragment of Bluescript KS+, isolated from HeLa cDNA library.

5A = human GABPβ2, Eco R1 fragment in Bluescript KS+, isolated from HeLa cDNA library.

F = human GABPβ3, Eco R1-Xho 1 fragment in Bluescript SK-, isolat d from human fetal brain cDNA library.

J = human GABPβ4, Eco R1- Xho 1 fragment in Bluescript SK-, isolated from human fetal brain cDNA library.

G = human GABPα, Eco R1- Xho 1 fragment in Bluescript SK-, from human fetal brain cDNA library.

* * * *

The entire contents of all documents cited herein are hereby incorporated by reference.

While various aspects of the invention have been described in some detail for purposes of clarity and understanding, one skilled in the art, from a reading of this disclosure will appreciate that various changes can be made in form and detail without departing from the true scope of the invention. One skilled in the art will also appreciate that the invention includes the human counterparts of the sequences specifically disclosed herein as well as the encoded amino acid sequences, and fragments thereof. The invention also relates to the deposited human sequences and fragments thereof as well as to the full length sequences of which the partial sequences form a part.

WHAT IS CLAIMED IS:

1. A DNA segment encoding a subunit of GA binding protein (GABP), or an epitope specific thereto, or a DNA fragment complementary to said DNA segment.

2. The DNA segment according to claim 1, wherein said GABP is human GABP.

3. The DNA segment according to claim 1 wherein said subunit is GABPα.

4. The DNA segment according to claim 1 wherein said subunit is GABPβ1.

5. The DNA segment according to claim 1 wherein said subunit is GABPβ2.

6. The DNA segment according to claim 3 wherein said subunit has the amino acid sequence shown in Figure 2A.

7. The DNA segment according to claim 4 wherein subunit has the amino acid sequence shown in Figure 2B.

8. The DNA segment according to claim 5 wherein said subunit has the amino acid sequence shown in Figure 2B.

9. A recombinant DNA molecule comprising:

i) said DNA segment according to claim 1; and

ii) a vector.

10. A host cell stably transformed with said recombinant DNA molecule according to claim 9.

11. The host cell according to claim 10 wherein said cell is a procaryotic cell.

12. The host cell according to claim 10 wherein said cell is a eucaryotic cell.

13. A method of producing a recombinant GABP subunit protein, or portion thereof defining at least an epitope specific thereto, comprising culturing said host cell according to claim 10 under conditions such that said segment is expressed and said protein thereby produced, and isolating said protein.

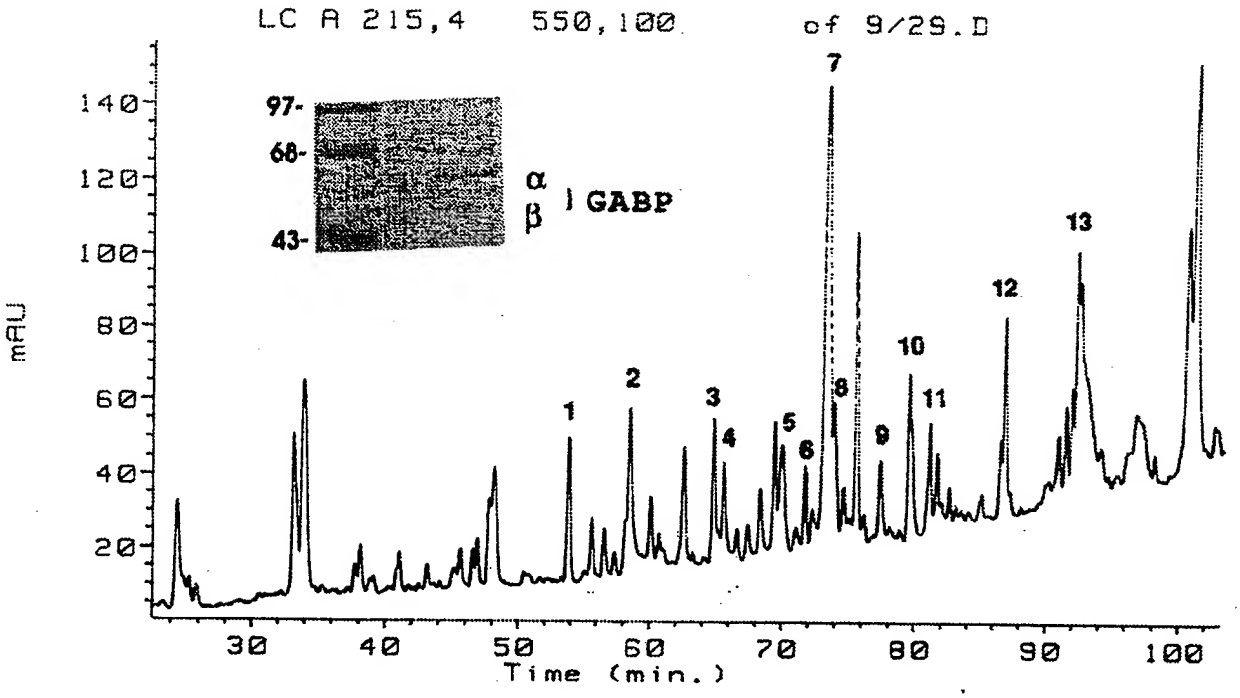

LC A 215,4 550,100 of 9/29.D
97-
68-
43-
α
β
GABP
mAU
140
120
100
80
60
40
20
30
40
50
60
70
80
90
100
Time (min.)
1
2
3
4
5
6
7
8
9
10
11
12
13

FIG. 1

# FIG. 2AI

**GABPα**

```
1    GCATGCGTATTCTGGCCTGTCCCCTAGTTCAAGCTCCCTTCGGAGTCAGCGTCCCGATCC
61   TTGGGGTTGGGGAGGTGGATGTAGGGTGGGGGTACAAAGGAGCGCGTAAAAACCAAAACC
121  AAAACAAAACAAAAAACCAAAGCGGCTGGGGGGCGTTTGTGCCTCGTGAGCCCCGTCGTC
181  GCGGGCGTTTCGGGCCGCGCAGCCTGCTGGGAGTTGTAGTTCCAGGCCCGGGGGTGCCTG
241  GGAGGCGGGAGAGGGGCTGGGGTTCCTTAGCGCCGATTCCACGGGAAGGGCCCCGGGACC
301  TTACCCTGCTACTCGCGGGAGCGGCGCGGCTCGCCCGGAGAGGCGCTGTACCTGGAGCCC
361  GCGCCGCCGCCGTTCCCAGCCGGCTGTGGAAGCGCGGGCTAGGGCCACGGGACGAATTCC
421  GGAGCTGGACTGAACTTTTGAACGTCTTCAACCATGACTAAGAGAGAAGCAGAAGAGCTG
                                    M  T  K  R  E  A  E  E  L          9
481  ATAGAAATTGAGATCGACGGGACTGAGAAAGCAGAGTGCACAGAAGAAAGCATTGTGGAA
     I  E  I  E  I  D  G  T  E  K  A  E  C  T  E  E  S  I  V  E      29
541  CAAACCTATACCCCAGCTGAATGTGTAAGCCAGGCCATAGACATCAATGAACCAATAGGC
     Q  T  Y  T  P  A  E  C  V  S  Q  A  I  D  I  N  E  P  I  G      49
601  AATTTAAAGAAACTACTAGAACCAAGACTGCAGTGTTCTTTGGATGCTCATGAAATTTGC
     N  L  K  K  L  L  E  P  R  L  Q  C  S  L  D  A  H  E  I  C      69
661  CTGCAAGATATTCAGCTGGATCCAGACCGAAGCTTGTTTGATCAAGGAGTGAAAACAGAT
     L  Q  D  I  Q  L  D  P  D  R  S  L  F  D  Q  G  V  K  T  D      89
721  GGGACTGTACAGCTTAGTGTACAGGTAATTTCTTACCAAGGAATGGAGCCAAAGTTGAAC
     G  T  V  Q  L  S  V  Q  V  I  S  Y  Q  G  M  E  P  K  L  N     109
781  ATTCTTGAAATTGTTAAGACTGCGGAAACGGTCGAGGTGGTCATCGATCCAGATGCCCAC
     I  L  E  I  V  K  T  A  E  T  V  E  V  V  I  D  P  D  A  H     129
```

# FIG. 2A2

```
841   CACGCGGAAGCAGAAGCGCATCTCGTTGAAGAAGCTCAAGTGATAACTCTTGACGGCACC
      H  A  E  A  E  A  H  L  V  E  E  A  Q  V  I  T  L  D  G  T      149

901   AAGCACATTACGACCATTTCAGACGAGACCTCGGAGCAGGTGACGAGATGGGCTGCTGCA
      K  H  I  T  T  I  S  D  E  T  S  E  Q  V  T  R  W  A  A  A      169

961   CTGGAAGGCTACAGAAAAGAGCAGGAGCGCCTTGGCATCCCCTATGATCCTATACGCTGG
      L  E  G  Y  R  K  E  Q  E  R  L  G  I  P  Y  D  P  I  R  W      189

1021  TCCACGGACCAAGTCCTGCATTGGGTGGTTTGGGTAATGAAGGAGTTCAGCATGACTGAT
      S  T  D  Q  V  L  H  W  V  V  W  V  M  K  E  F  S  M  T  D      209

1081  ATAGACCTCACCACACTCAACATTTCGGGAAGAGAATTATGTAGTCTCAACCAAGAAGAT
      I  D  L  T  T  L  N  I  S  G  R  E  L  C  S  L  N  Q  E  D      229

1141  TTTTTTCAGCGGGTCCCTCGGGGAGAAATTCTTTGGAGTCATCTGGAGCTTCTTCGAAAA
      F  F  Q  R  V  P  R  G  E  I  L  W  S  H  L  E  L  L  R  K      249

1201  TATGTTTTGGCCAGCCAAGAGCAACAGATGAATGAGATAGTTACCATTGACCAGCCTGTG
      Y  V  L  A  S  Q  E  Q  Q  M  N  E  I  V  T  I  D  Q  P  V      269

1261  CAGATTATTCCAGCCTCAGTGCCTCCCGCTACACCGACTACGATTAAAGTTATAAACAGC
      Q  I  I  P  A  S  V  P  P  A  T  P  T  T  I  K  V  I  N  S      289

1321  AGTGCAAAAGCAGCTAAAGTGCAACGGTCCCCAAGGATTTCAGGAGAAGACAGAAGTTCA
      S  A  K  A  A  K  V  Q  R  S  P  R  I  S  G  E  D  R  S  S      309

1381  CCGGGGAACAGAACAGGAAACAATGGTCAGATCCAACTATGGCAGTTTTTGCTAGAACTT
      P  G  N  R  T  G  N  N  G  Q  I  Q  L  W  Q  F  L  L  E  L      329
```

# FIG. 2A3

```
1441  CTTACTGACAAGGATGCTCGAGACTGTATTTCTTGGGTTGGTGATGAAGGTGAATTTAAG
      L  T  D  K  D  A  R  D  C  I  S  W  V  G  D  E  G  E  F  K      349

1501  CTAAATCAGCCTGAGTTGGTTGCGCAAAAATGGGGACAACGTAAGAACAAGCCTACCATG
      L  N  Q  P  E  L  V  A  Q  K  W  G  Q  R  K  N  K  P  T  M      369

1561  AACTATGAGAAACTTAGCCGTGCATTACGGTATTATTATGATGGGGACATGATTTGTAAA
      N  Y  E  K  L  S  R  A  L  R  Y  Y  Y  D  G  D  M  I  C  K      389

1621  GTTCAAGGCAAGAGATTTGTGTACAAATTTGTTTGTGACTTGAAGACTCTTATTGGATAC
      V  Q  G  K  R  F  V  Y  K  F  V  C  D  L  K  T  L  I  G  Y      409

1681  AGTGCAGCAGAACTGAACCGTCTGGTCATAGAGTGTGAACAGAAGAAACTGGCACGGATG
      S  A  A  E  L  N  R  L  V  I  E  C  E  Q  K  K  L  A  R  M      429

1741  CAGCTGCATGGGATTGCCCAGCCAGTCACGGCAGTAGCACTGGCAGCCACCTCTCTACAG
      Q  L  H  G  I  A  Q  P  V  T  A  V  A  L  A  A  T  S  L  Q      449

1801  GCAGACAAAGAGATTTGAGACCTAGGACCTCCTGGGGAGTCTTAAGGTTTTTCTTAAATA
      A  D  K  E  I                                                   454

1861  TTTAGAGCAAGCTTTGCTCTAACCTTTATTACTGAATTTGAATCGTATTTCTAGAGTGTA
1921  CAATCTGATGCATGATTTTTTTTATAAATATTTCATACTCTTGTGAAAAAAAAAAAAAAA
1981  AAAAA  1985
```

GABP$\beta$1

# FIG. 2BI

```
   1  GGAAAATAGCGCCCTGTGCAGCGAAGCGCTCTGTGTGTGGCGGGGCCGCGTCTGCCCTGA
  61  CGGCTCCGAGGCGCCGCGGTCCCCGCACCTCCTCCCGCTGCTCCCCCGCCGCCGCTCCCG
 121  AAGCTTTTCCAGATGTCCCTGGTAGATTTGGGGAAGAAGCTTTTAGAAGCGGCACGAGCC
                M  S  L  V  D  L  G  K  K  L  L  E  A  A  R  A      16

 181  GGTCAAGATGATGAAGTTCGCATTTTGATGGCAAATGGAGCTCCTTTTACTACAGACTGG
      G  Q  D  D  E  V  R  I  L  M  A  N  G  A  P  F  T  T  D  W      36

 241  TTGGGAACTTCTCCACTTCATCTGGCCGCACAGTATGGGCATTTCTCTACCACAGAGGTT
      L  G  T  S  P  L  H  L  A  A  Q  Y  G  H  F  S  T  T  E  V      56

 301  CTTCTCCGAGCCGGTGTAAGTAGAGATGCCAGGACCAAAGTGGACCGGACACCACTGCAC
      L  L  R  A  G  V  S  R  D  A  R  T  K  V  D  R  T  P  L  H      76

 361  ATGGCGGCTTCTGAGGGCCATGCCAACATAGTAGAAGTTTTGCTTAAGCATGGTGCTGAC
      M  A  A  S  E  G  H  A  N  I  V  E  V  L  L  K  H  G  A  D      96

 421  GTCAATGCAAAGGATATGTTAAAGATGACAGCTCTGCATTGGGCAACAGAACATAATCAT
      V  N  A  K  D  M  L  K  M  T  A  L  H  W  A  T  E  H  N  H     116

 481  CAAGAGGTGGTGGAGCTTTTAATCAAATATGGTGCTGATGTACACACGCAGAGTAAATTT
      Q  E  V  V  E  L  L  I  K  Y  G  A  D  V  H  T  Q  S  K  F     136

 541  TGTAAAACTGCATTTGATATTTCAATAGACAATGGAAATGAAGATTTAGCAGAGATATTA
      C  K  T  A  F  D  I  S  I  D  N  G  N  E  D  L  A  E  I  L     156

 601  CAGATTGCTATGCAGAACCAAATCAACACCAACCCGGAGAGTCCTGACACTGTGACAATA
      Q  I  A  M  Q  N  Q  I  N  T  N  P  E  S  P  D  T  V  T  I     176
```

# FIG. 2B2

```
661   CACGCTGCCACACCACAGTTCATCATTGGACCCGGAGGGGTGGTGAACCTCACAGATGAA
      H  A  A  T  P  Q  F  I  I  G  P  G  G  V  V  N  L  T  D  E      196

721   ACAGGAGTATCTGCTGTTCAGTTTGGAAACTCCTCTACGTCAGTATTAGCTACGTTAGCT
      T  G  V  S  A  V  Q  F  G  N  S  S  T  S  V  L  A  T  L  A      216

781   GCCTTAGCTGAAGCTTCTGCCCCATTGTCCAATTCTTCAGAAACTCCAGTGGCCACAGAG
      A  L  A  E  A  S  A  P  L  S  N  S  S  E  T  P  V  A  T  E      236

841   GAAGTGGTTACCGCAGAATCTGTGGATGGTGCAATTCAGCAAGTAGTTAGCTCAGGGGGT
      E  V  V  T  A  E  S  V  D  G  A  I  Q  Q  V  V  S  S  G  G      256

901   CAGCAAGTCATCACGATAGTTACAGATGGAATCCAGCTGGGGAATTTGCACTCCATACCA
      Q  Q  V  I  T  I  V  T  D  G  I  Q  L  G  N  L  H  S  I  P      276

961   ACCAGTGGGATGGGTCAGCCCATCATTGTGACGATGCCCGATGGACAGCAAGTATTGACA
      T  S  G  M  G  Q  P  I  I  V  T  M  P  D  G  Q  Q  V  L  T      296

1021  GTACCAGCAACAGACATTGCTGAAGAAACTGTCATCAGTGAAGAGCCACCAGCTAAGAGA
      V  P  A  T  D  I  A  E  E  T  V  I  S  E  E  P  P  A  K  R      316

1081  CAGTGTATGGAAATAATTGAGAGCCGGGTGGAATGTGCAGAAATTGAAGAGAGAAGCG
      Q  C  M  E  I  I  E  S  R  V  E  C  A  E  I  E  E  R  E  A      336

1141  CTTCAGAAACAGCTGGATGAAGCCAACCGAGAGGCCCAGAAATACCGACAGCAGCTGCTT
      L  Q  K  Q  L  D  E  A  N  R  E  A  Q  K  Y  R  Q  Q  L  L      356
```

# FIG. 2B3

```
1201    AAGAAGGAGCAGGAGGCAGAGGCCTACAGGCAGAAGCTGGAGGCCATGACACGCATCCAG
        K _K_ E _Q__E_ A _E_ A _Y__R_ Q _K_ L _E__A_ M _T_ R _I_ Q      376

1261    ACCAACAAAGAAGCCGTTTAGCTGCCATGAACACCAGTTTGCTTTTACCTTTTGTCCAGA
        T _N_ K_ E _A_ V                                               382

1321    AAGAATATAGTCTTGAACTGTACACAGTAAGGACACAGCCATGGGAATACCGAATAATAG
1381    AAAATACTACAGCTTGATAACGGGACTTAAGCCATGAGCTCTCAATTCTTGTAATATAAA
1441    ACTTTAGAACTTGTGAAATGTATTTAAAACTGAATTCTGTAAATAGTTTTTTGGTTGTTT
1501    CTTTTTTTACTGTTATAAATGAGTTGATGAAGCTTGTTGCTGAGATCCAGAGGCCACAGC
1561    AAGCCACTGTTACGGGGAATTCTTTTTGATTTCGGTGCAGACTTAATTTCTCAGAAACAG
1621    AAGTTGTAAGCGTGGTTGTTGTCTCTTAGACCAAATGTGGAACAATTATGGTGGACTGAT
1681    GCTGGAAATACTCTGCAAAGAGATTTCGCCCAGGAACTTTTGTACAGCTTTTAAGTTGTC
1741    TCAGGTTCTCTGAAAACATTTTTTAGGAAGCAGAAATTTTATATTTGTTCTATTTCAGCT
1801    ACATGCAAGTAGATTTACATGTATATGAAGCAAATATTTTTAAAATTTTGTTTGTACATA
1861    TCCTGCATGTTTTATGATTTTAAAACACATCCCTTGCTTAGGTATGTCTCCTACAAAGAC
1921    GATGAAAGTTACCAAGAAAAAAACCAAGGCACCACTGTCGTTCTGCCGGTCGTTGAGAGT
1981    GAGTAAGCTCACAGCTGGTTTTACTGGCGACCATGACAATGTTCTTGGAGCAGCTTGCAA
2041    GAGAACTCGAATCTGCAAAATATATGCTTTTCTGACAATTGTATTGTGAGGATGTGAGAC
2101    TGCTCTGTGCCTTCCTTCCATCATGACTTCCACACGAGCACTTTGAGGCGCTGGATTTTA
2161    AGATAATGTTTTTGGCAAACTCAGTGCATTTGGGTTTCTGAAATATTTCATGGACTTACT
2221    CCCCTCCCCTAAGAAATTAAATTATTCTTATGGAAATTGCTTTTGTATGCAGAGCAGGAC
2281    CTTTTGTACTCAGTGAGGCTATGGGTGTGTCTAACGTAACTTTCGGTATTCCCTGTGAAC
2341    GCGCTCCGGGCCTGTGCTGCGACTCTCTCACAATATTGTCAATTCTGAAATGTATGGACA
2401    TTTGACAGACTATTCCATTCTCTTTTGTAAAATGTAACTTCAAGCTCTTTATTATTTTGT
2461    TTTGTTCTATTGCTTTATGAAAACGTTTGCCCCTCCCCCACCCCAGGCTCCTCTAGCTTA
2521    TCTGTATAAAGAACCAGTTACCTGCGAGACTGTACTCTGGCCTCATTTTCTATTTTACAA
2581    TTAAATAACTTTCCACATAAAAAAAAAAAAAAAAA  2616
```

FIG. 2B4

GABP$\beta$2

```
1131    GTAAGGAGCCTTATACCCGGTGTGTTTTGCTGCAGCCATCCAAAATAAATTTTGTTTCAT
        V  R  S  L  I  P  G  V  F  C  C  S  H  P  K                  349

1191    TATTATTATTATTATCACTATTATTATATTAATTTGTTGTTTATATTTATTACTGGCCAT
1251    GTTTTGATAAAGAATGAAAGTGAATAGTGTCATCATTTTGTTTTGGCCTTAAACATATGG
1311    CAAGCCCTCAATAAATAAATATTGAATGAAAAAAAAAAAAAAAAA  1356
```

FIG. 3B

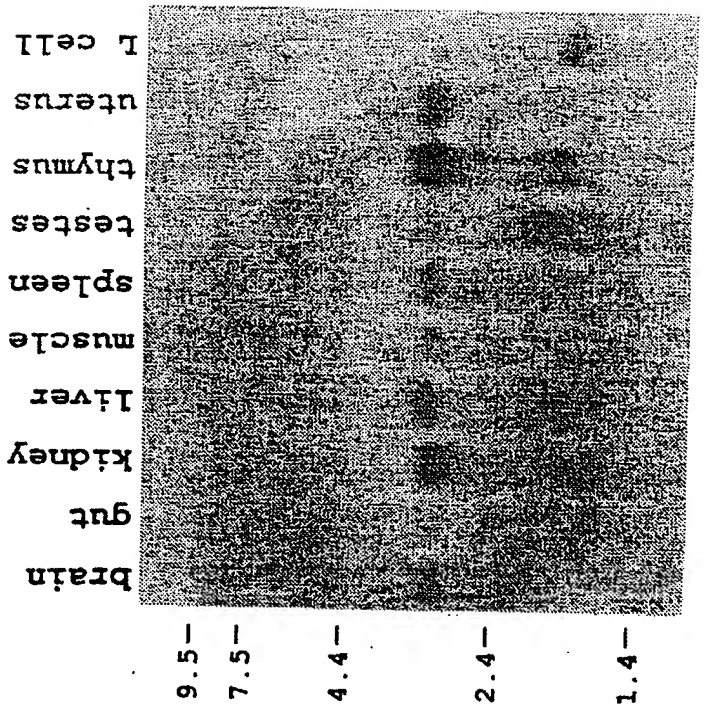

brain
gut
kidney
liver
muscle
spleen
testes
thymus
uterus
L cell
9.5
7.5
4.4
2.4
1.4

FIG. 3A

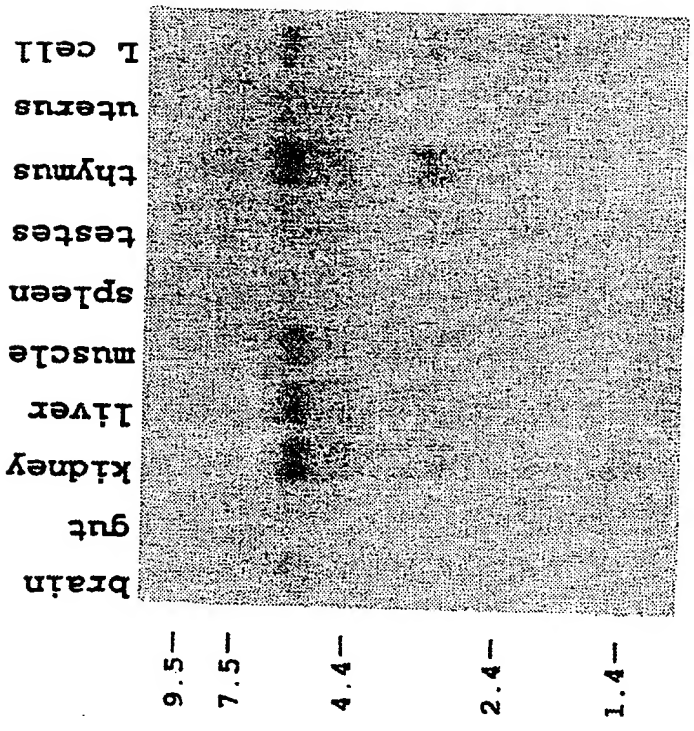

brain
gut
kidney
liver
muscle
spleen
testes
thymus
uterus
L cell
9.5
7.5
4.4
2.4
1.4

FIG. 4A

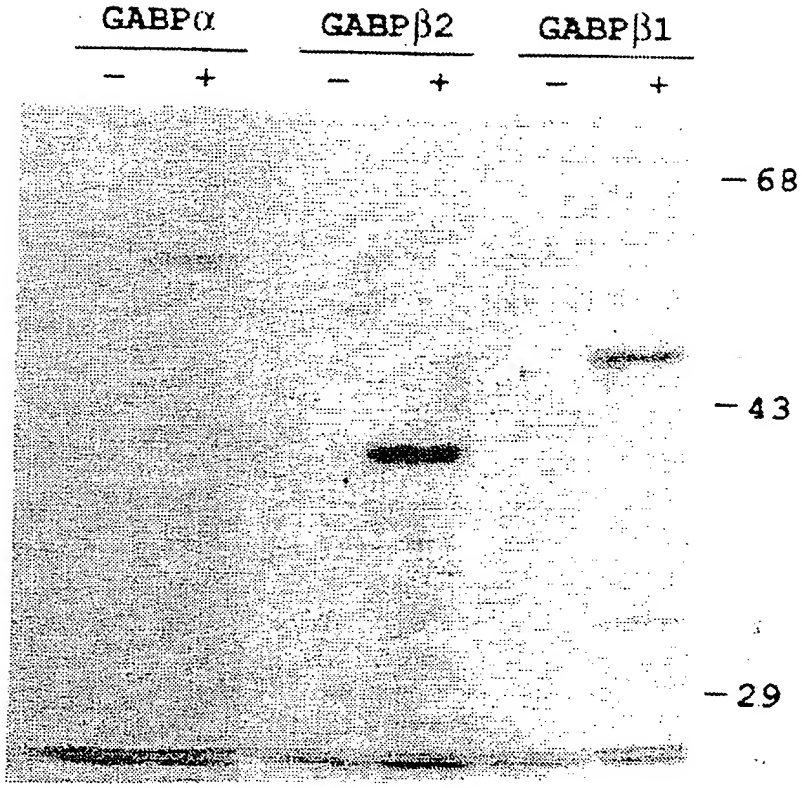

GABPα
GABPβ2
GABPβ1
− + − + − +
68
43
29

FIG. 4B

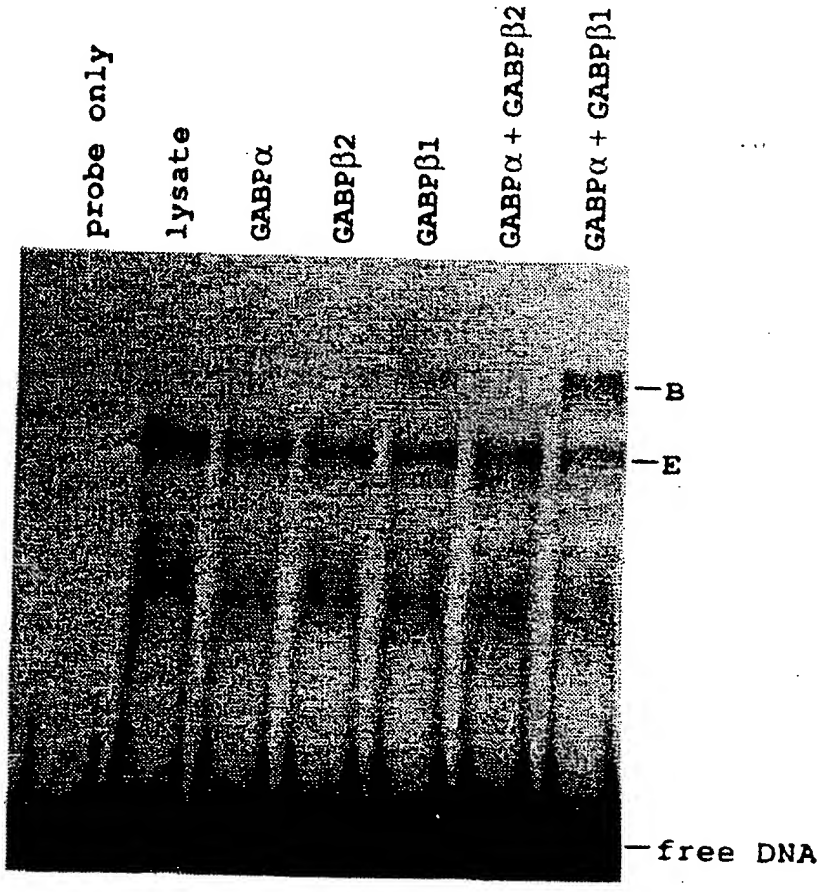
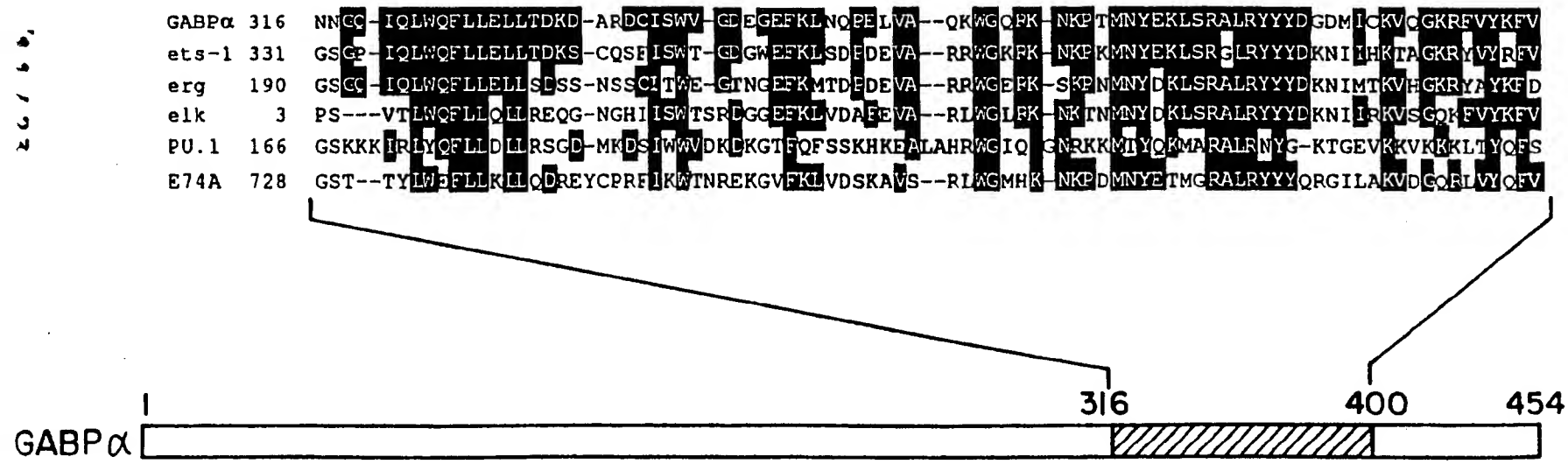
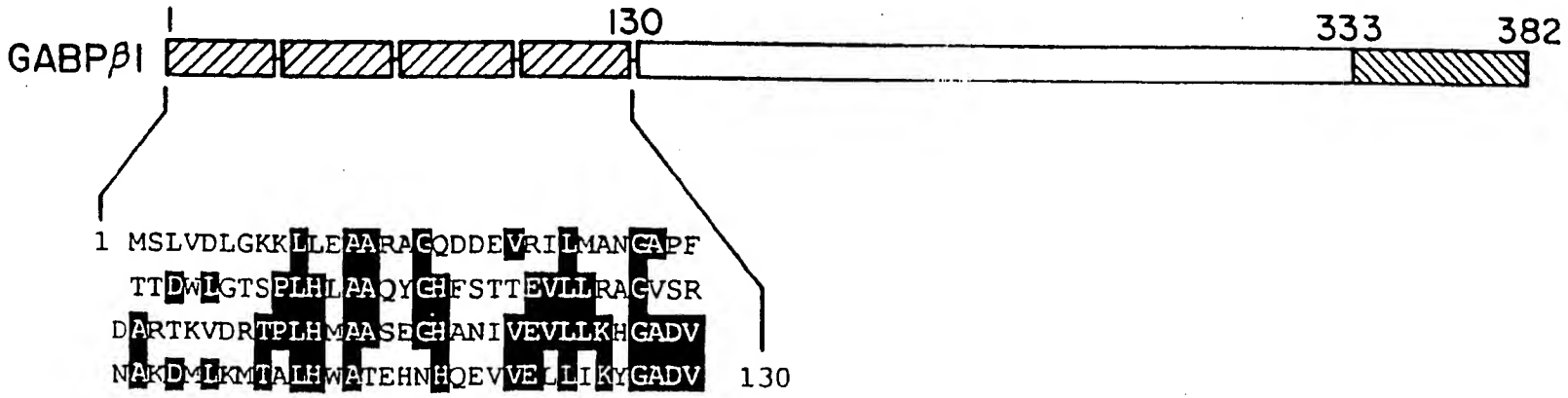

probe only
lysate
GABPα
GABPβ2
GABPβ1
GABPα + GABPβ2
GABPα + GABPβ1
B
E
free DNA

FIG. 5


GABPα 316
ets-1 331
erg 190
elk 3
PU.1 166
E74A 728
1
316
400
454
GABPα




| | |
|---|---|
| GABPβ | -A-D-L--TPLH-AA---H---VEVLLK-GADV |
| cdc10/SWI4,6 | N-QD--G-TPLHWAA--AN--LV--L---GANL |
| Notch | NA-D--G-TPLHLAAR---------LL---A-- |
| glp1 | ---D--GRTALHLAA-------V--L------- |
| lin12 | ---D--GRTALH-AA-N-----V-YL----A-- |
| ankyrin | -D----G-TPLH-AA--GH---V--LL--GA-- |
| NFκB | ---D--G-T-LHLA-------L---LL--GAD- |
| fem1 | ------G-TPL--AA--GH---VVKLLE-G-D- |
| bcl3 | -A-N--G-TPLH-AV----L--VRLLL--GA-- |
| overall consensus: | ---D--G-TPLH-AA-------V--LL--GA-- |

FIG. 5 con't.

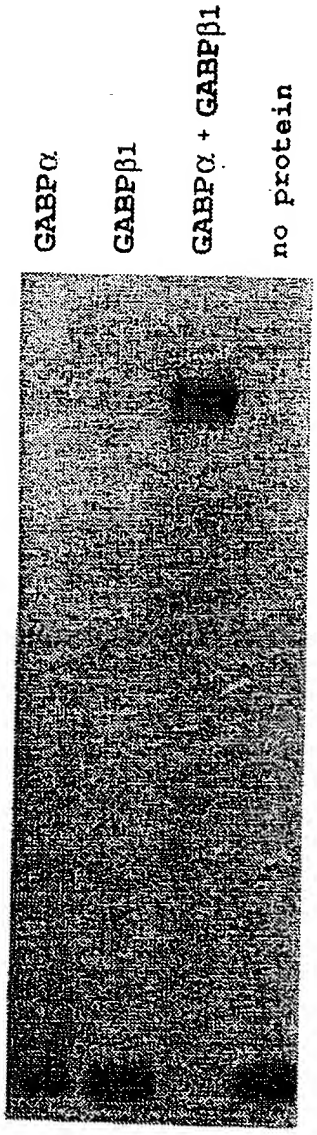

GABPα
GABPβ1
GABPα + GABPβ1
no protein

FIG. 6A

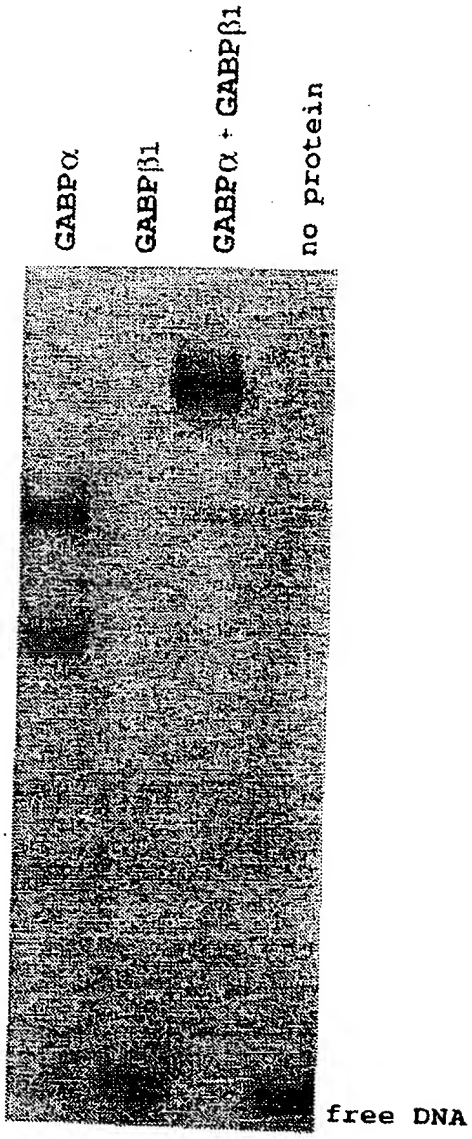

GABPα
GABPβ1
GABPα + GABPβ1
no protein
free DNA

FIG. 6B

FIG. 7A1 FIG. 7A2

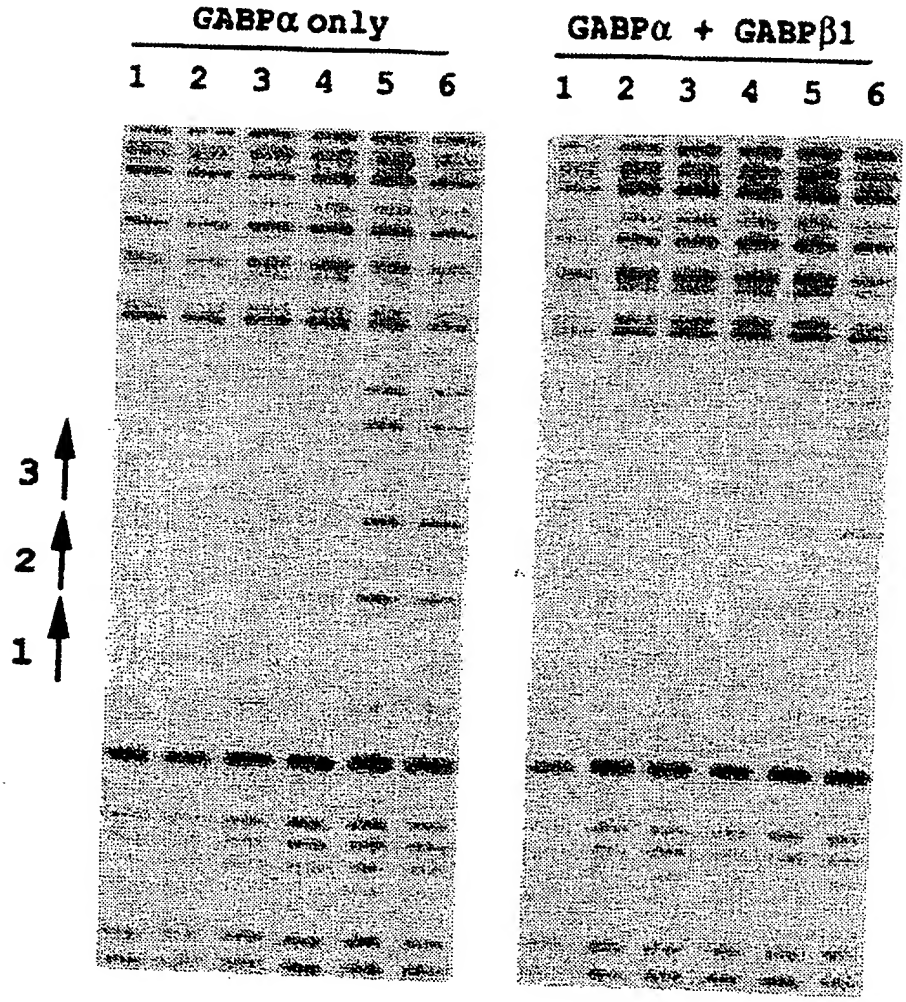


FIG. 7B1

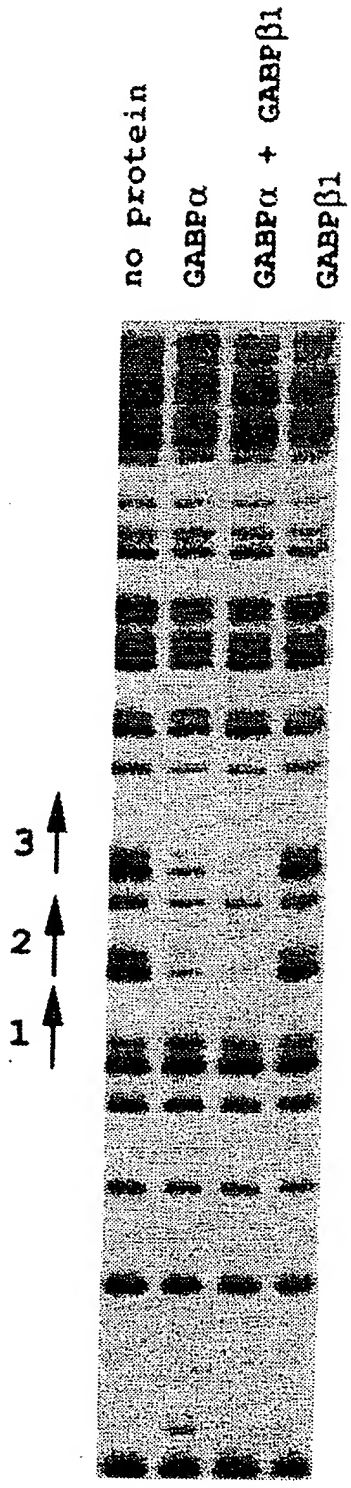
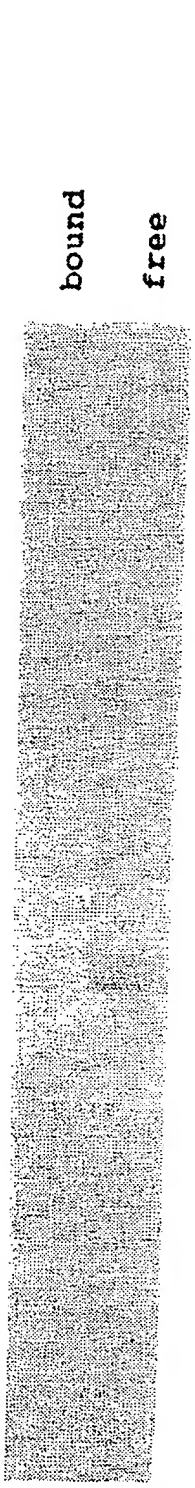



FIG. 7B2




TGCGGAACGGAAGCGGAAACCG

1 2 3

FIG. 8

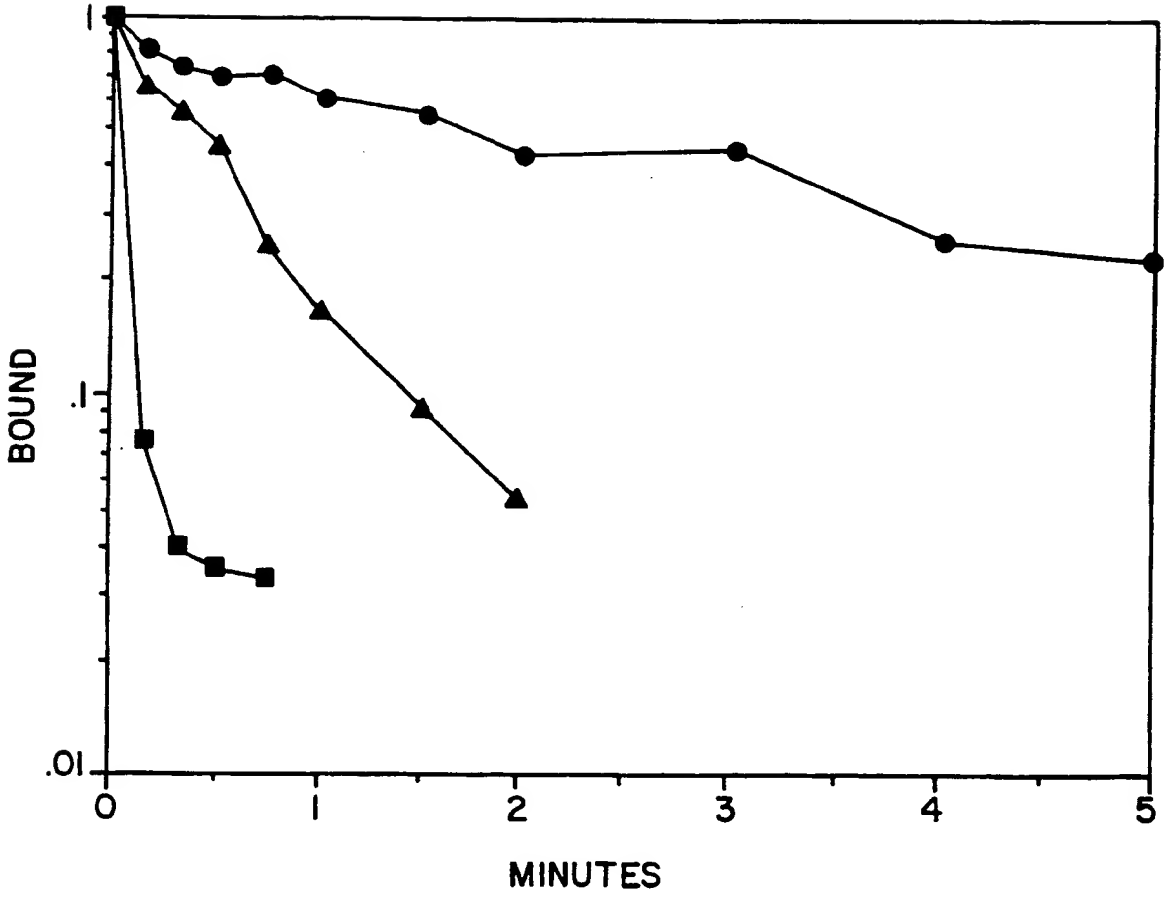


FIG. 9

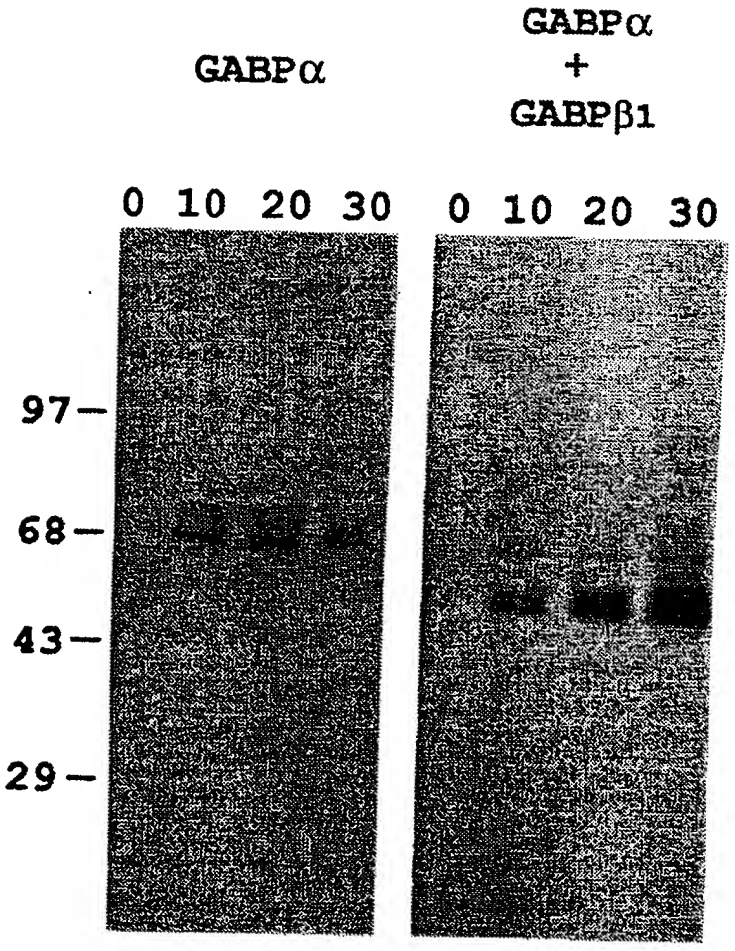


FIG. 10

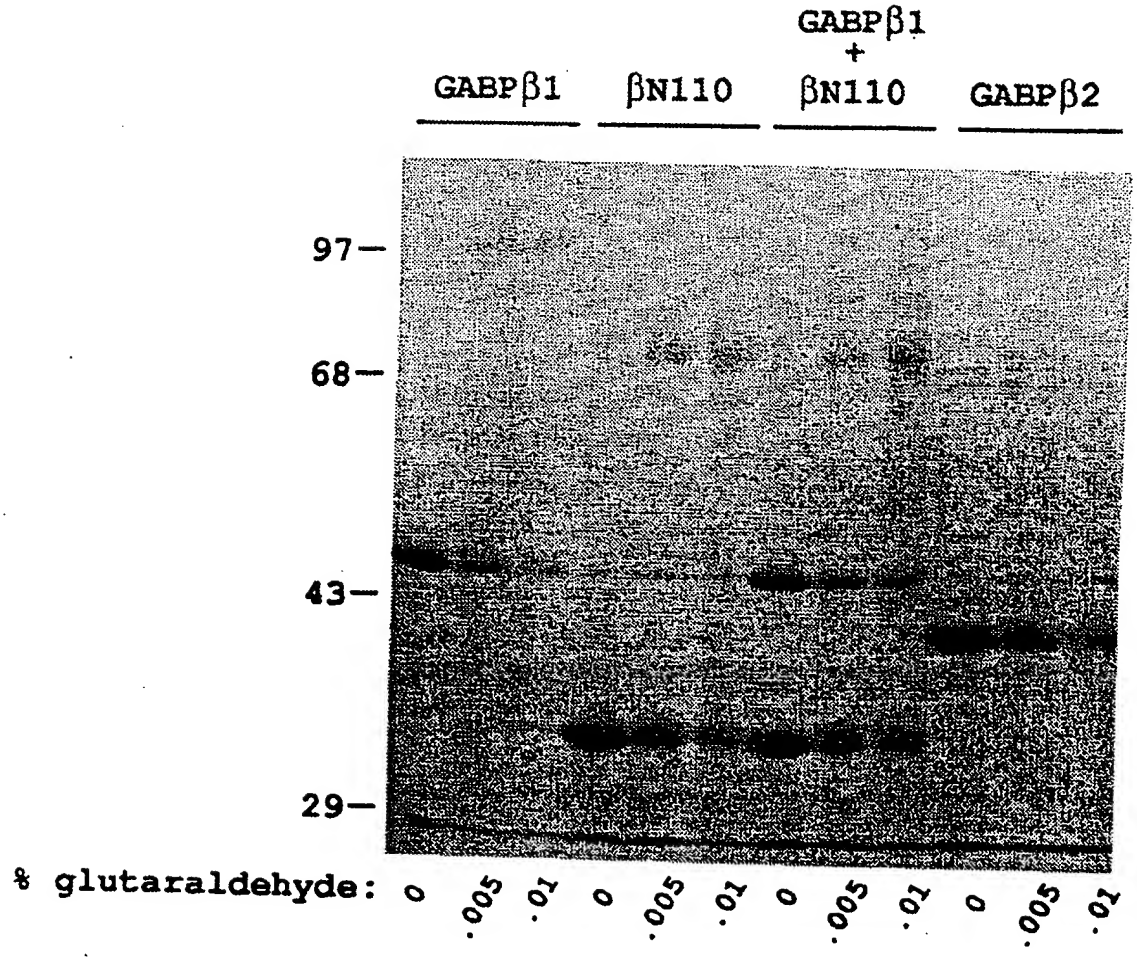

GABPβ1
+
GABPβ1
βN110
βN110
GABPβ2
97—
68—
43—
29—
% glutaraldehyde:
0
.005
.01
0
.005
.01
0
.005
.01
0
.005
.01

FIG. 11A

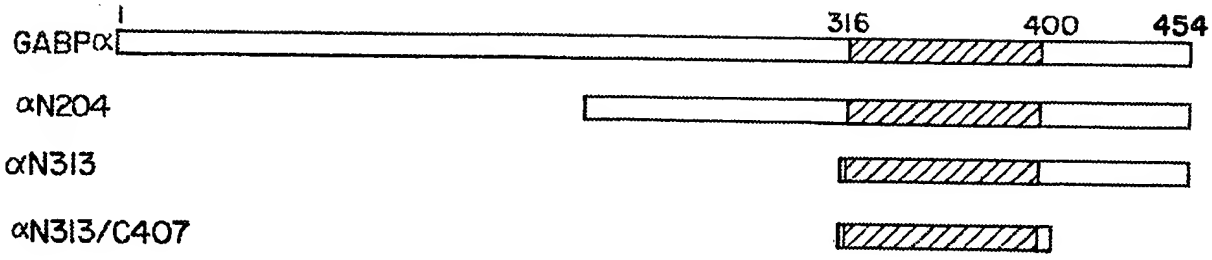
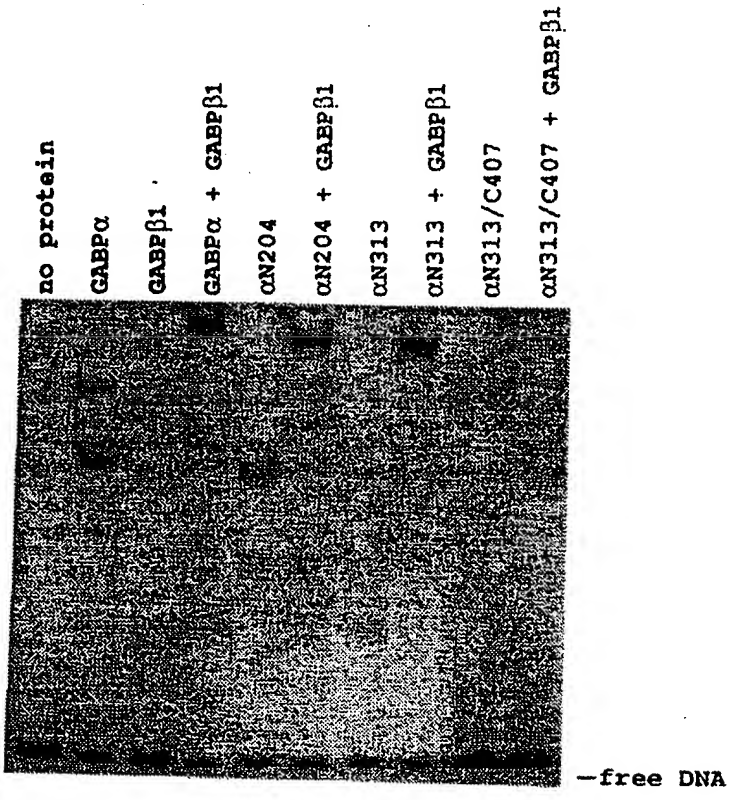



FIG. 11B



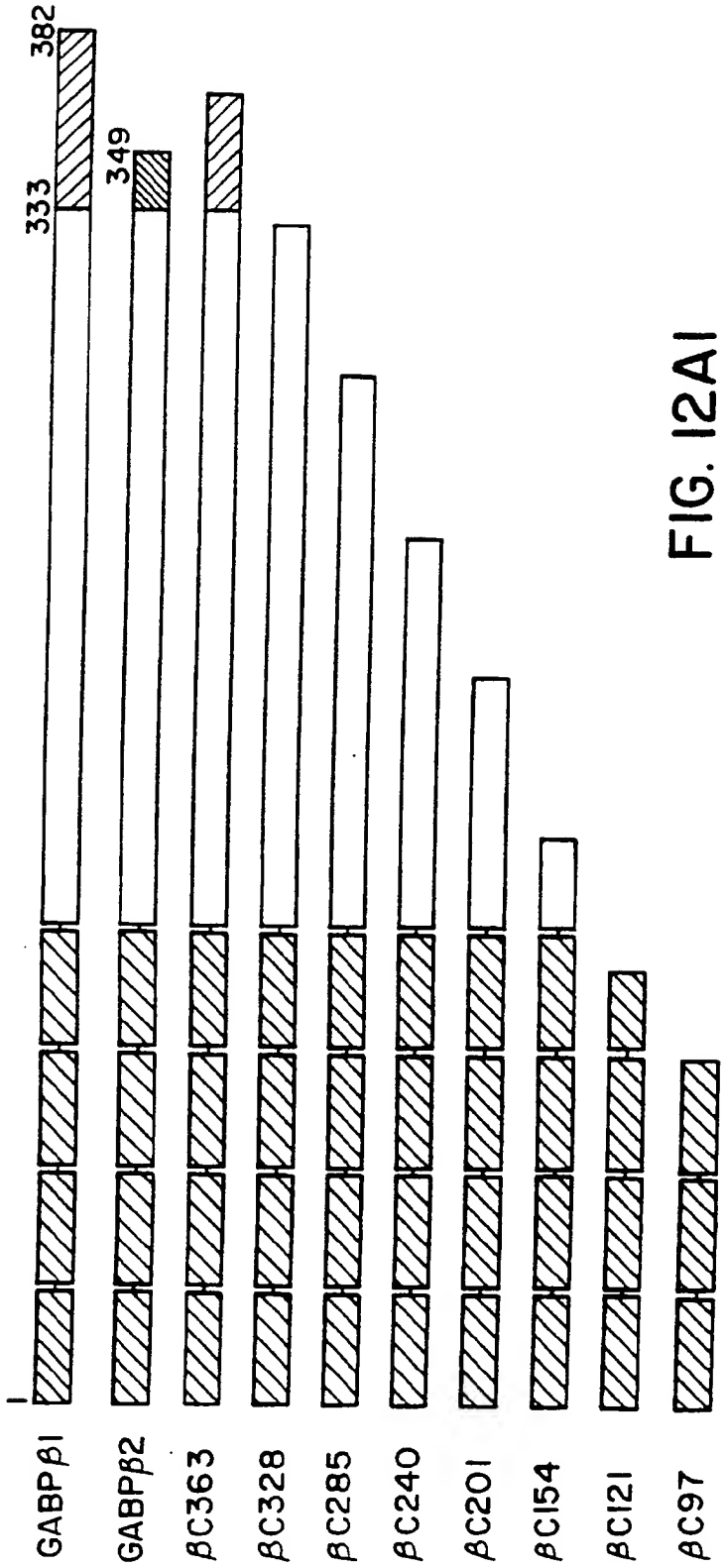

382
349
333
1
GABPβ1
GABPβ2
βC363
βC328
βC285
βC240
βC201
βC154
βC121
βC97

FIG. 12A1

FIG. 12A3

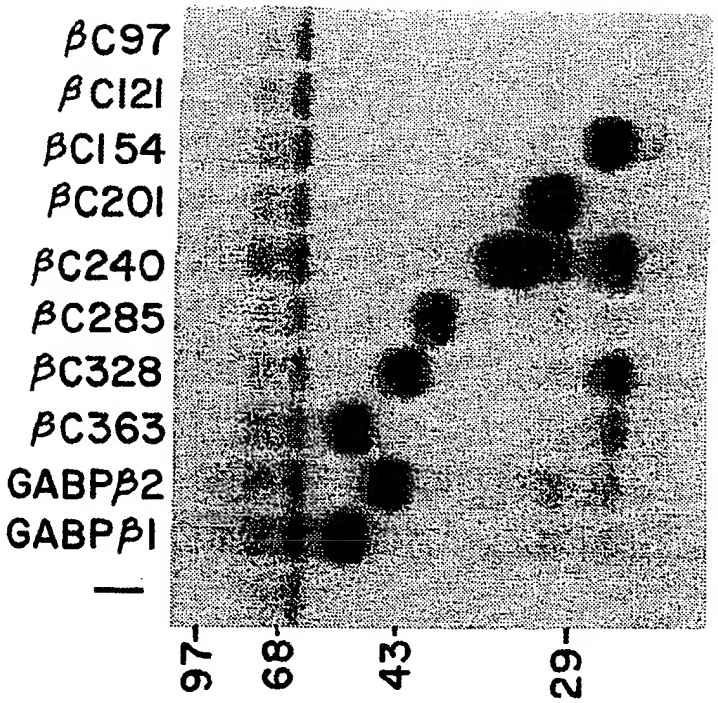



FIG. 12A2

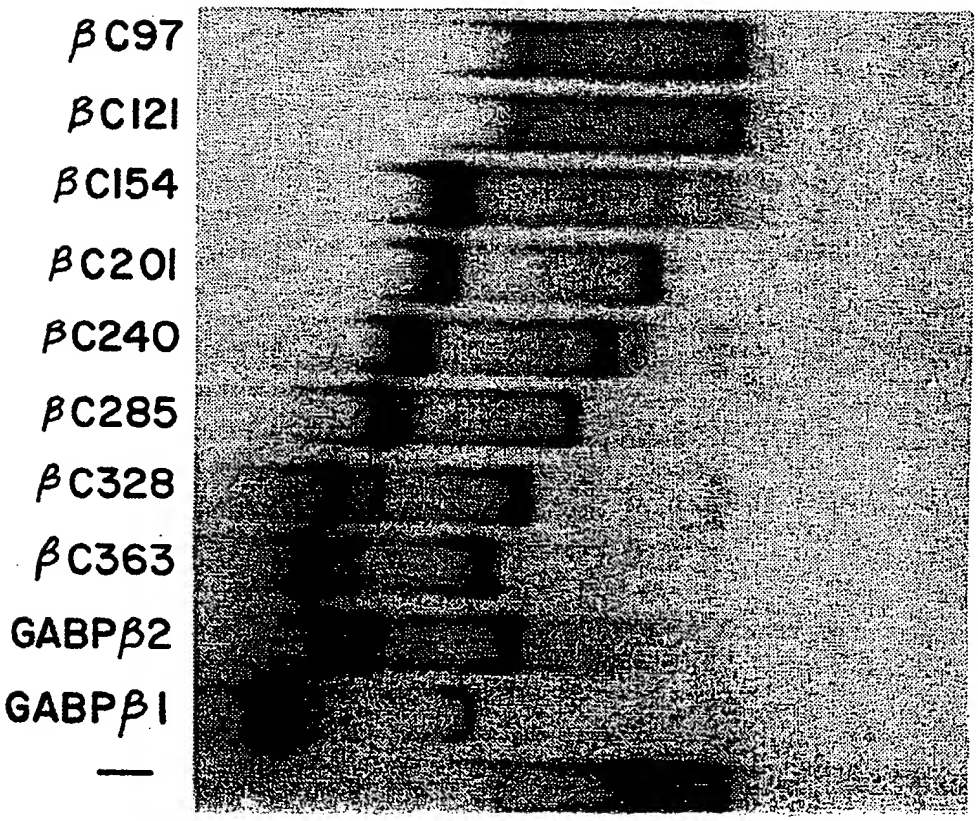


FIG. 12B1

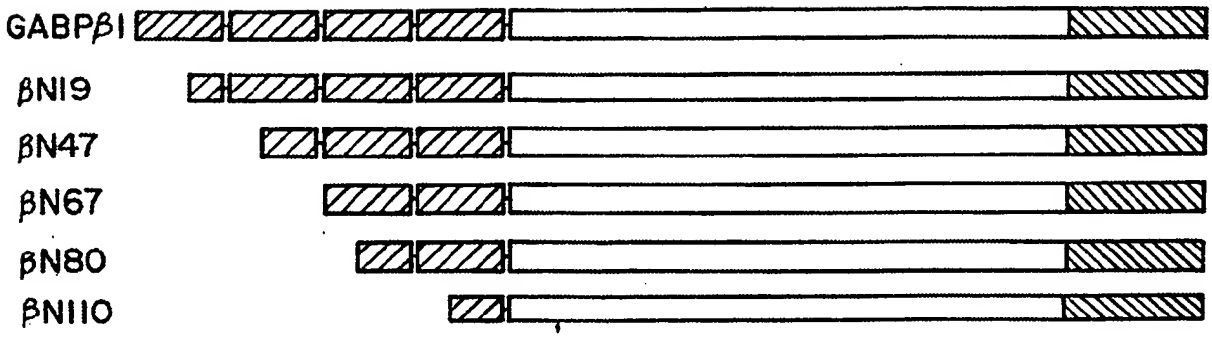



FIG. 12B2

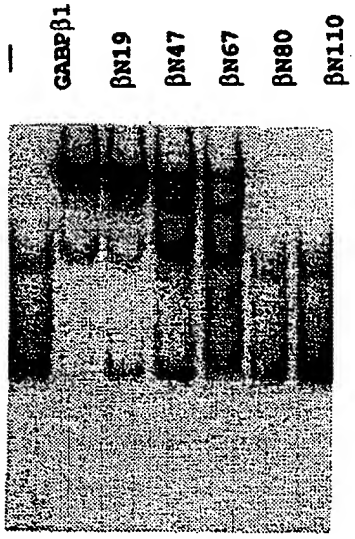



FIG. 12B3

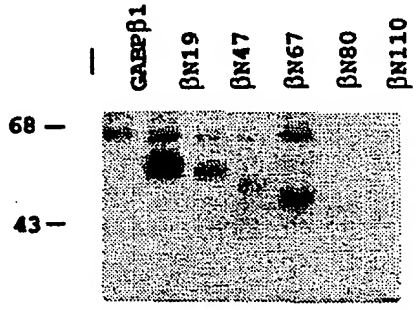
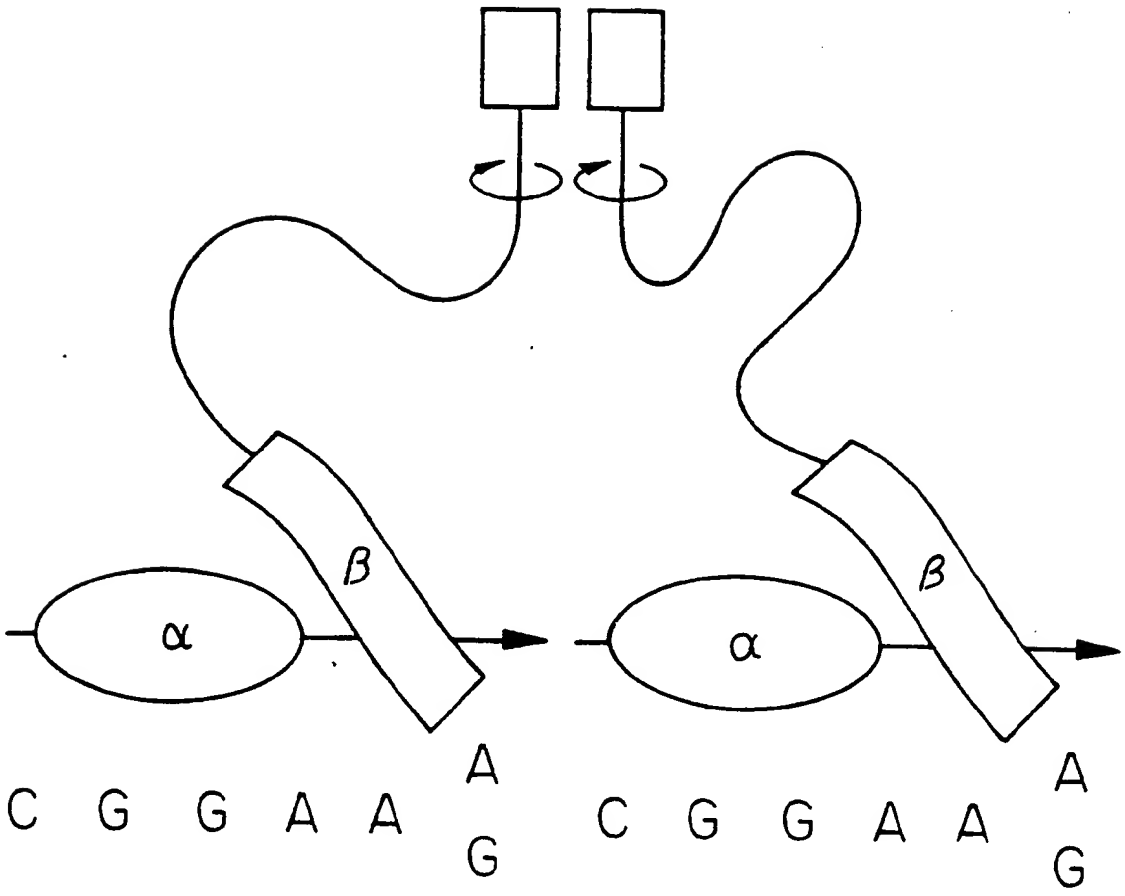


FIG. 13



# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/06748

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12N 15/00, 15/12, 15/63, 15/70, 15/74, 15/79
US CL :435/69.1, 172.3, 252.3, 320.1; 536/27; 935/10, 11, 27, 70, 72
According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 172.3, 252.3, 320.1; 536/27; 935/10, 11, 27, 70, 72

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | GENES AND DEVELOPMENT, Volume 3, issued 1989, K.L. LaMarco et al, "Purification of a set of cellular polypeptides that bind to the purine-rich cis-regulatory elemeent of herpes simplex virus immediate early genes", pages 1372-1383, see entire document. | 1-13 |
| Y | CELL, Volume 62, issued 07 September 1990, S. Ghosh et al, "Cloning of the p50 binding subunit of NF-kB: Homology to rel and dorsal", pages 1019-1029, see entire document. | 1-13 |
| Y | THE EMBO JOURNAL, Volume 7, No. 3, issued 1988, K.E. Boulukos et al, "Identification of an evolutionarily conserved cellular ets-2 gene (c-ets-2) encoding nuclear proteins related to the products of the c-ets proto-oncogene", pages 697-705, see entire document. | 1-3, 6 and 9-13 |

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of parucular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 14 NOVEMBER 1992 | [illegible] |

## INTERNATIONAL SEARCH REPORT

Inter onal application No.
PCT/US92/06748

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 265, No. 14, issued 15 May 1990, J.K. Nyborg et al, "Interaction of cellular proteins with the human T-cell leukemia virus type I transcriptional control region", pages 8230-8236, see entire document. | 1-3 and 6 |
| Y | SCIENCE, Volume 249, issued September 1990, C. Coffman et al, "Xotch, the Xenopus homolog of Drosophila Notch", pages 1438-1441, see entire document. | 4, 5 and 7-12 |
| Y | ONCOGENE RESEARCH, Volume 3, issued 1988, E.S.P. Reddy et al, "Structure, expression and alternative splicing of the human c-ets-1 proto-oncogene", pages 239-246, see pages 240-245. | 4, 5 and 7-13 |
| Y | WO, A, 90/05745 (Habener et al) 31 May 1990, pages 14-27. | 9-13 |
| Y | CELL, Volume 61, No. 1, issued 6 April 1990, M.J. Klemsz et al, "The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene", pages 113-124, see entire document. | 1-11 |
| Y | ONCOGENE, Volume 5, issued 1990, D. Ray et al, "The human homologue of the putative proto-oncogene Spi-1: Characterization and expression in tumors", pages 663-668, see entire document. | 1-11 |
| Y | JOURNAL OF VIROLOGY, Volume 65, No. 1, issued January 1991, R. Paul et al, "The Sfpi-1 proviral integration site of Friend erythroleukemia encodes the ets-related transcription factor Pu.1", pages 464-467, see entire document. | 1-11 |
| A | WO, A, 91/07423 (Liao et al) 30 May 1991, pages 12-39. | 8-13 |
| A | WO, A, 89/09777 (Sukhatme) 19 October 1989, pages 9-22. | 9-11 and 13 |
| A | GENES AND DEVELOPMENT, Volume 2, issued 1988, S.J. Triezenberg et al, "Evidence of DNA:protein interactions that mediate HSV-1 immediate early gene activation by VP16", pages 730-742, see pages 735-741. | 1, 9-11 |
| A | CELL, Volume 60, issued 23 March 1990, A.M. Spence et al, "The product of fem-1, a nematode sex-determining gene, contains a motif found in cell cycle control proteins and receptors for cell-cell interactions", pages 981-990, see pages 982-988. | 1, 9-12 |
| A | NATURE, Volume 348, issued 01 November 1990, V. Bours et al, "Cloning of a mitogen-inducible gene encoding a kB DNA-binding protein with homology to the rel oncogene and to cell-cycle motifs", pages 76-80, see entire document. | 1, 9-11 |
| A | MOLECULAR AND CELLULAR BIOLOGY, Volume 8, No. 5, O.M. Andrisani et al, "Three sequence-specific DNA-protein complexes are formed with the same promoter element essential for expression of the rat somatostatin gene", pages 1947-1956, see pages 1948-1955. | 1, 9-12 |
| A | SCIENCE, Volume 245, issued 28 July 1989, P.J. Mitchell et al, "Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins", pages 371-378, see entire document. | 1 |
| A | THE JOURNAL OF BIOLOGICAL CHEMISTRY, Volume 265, No. 36, issued 25 December 1990, J. Ozer et al, "Isolation and characterization of a cDNA clone for the CCAAT transcription factor $EFI_A$ reveals a novel structural motif", pages 22143-22152, see pages 22144-22152. | 1, 9-11 |
| A | GENES AND DEVELOPMENT, Volume 4, issued 1990, C.-M. Fan et al, "A DNA- | 1, 9-11 |

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/06748

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| A | PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, USA, Volume 88, issued February 1991, R. Meyer et al, "Cloning of the DNA-binding subunit of human nuclear factor kB: The level of its mRNA is strongly regulated by phorbol ester or tumor necrosis factor $\alpha$", pages 966-970, see entire document. | 1-2 and 9-11 |
| A,P | EP, A, 0,449,170 (Hoppe-Seyler et al) issued 02 October 1991, pages 4-7. | 1, 9-11 |
| A | GENES AND DEVELOPMENT, Volume 4, issued 1990, C.V. Gunther et al, "Sequence-specific DNA binding of the proto-oncoprotein ets-1 defines a transcriptional activator sequence within the long terminal repeat of the Moloney murine sarcoma virus", pages 667-679. | 1, 9-11 |
| Y | JOURNAL OF VIROLOGY, Volume 62, No. 9, issued September 1988, D. LePrince et al, "Alternative splicing with the chicken c-ets-1 locus: Implications for transduction within the E26 retrovirus of the c-ets proto-oncogene", pages 3233-3241, see especially pages 3235-3240 and Figure 7. | 4-5, 7-8 |

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/06748

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

STN/Chemical Abstracts; DIALOG/Biosis, Derwent Biotechnology Abstracts, Current Biotechnology Abstracts, Agricola, Medline; EMBL-NEW 7; UEMBL 31-72; GenBank-NEW 7; GenBank 72; N-GeneSeq 7, PIR 32; SwissProt 22; A-GeneSeq 7

83.5% identity in 746 nt overlap; score: 2027

```
Murine Rin2 540       550       560       570       580       590
_       AACAGACGAAGATGTTTTTGGAAGCAATGCCTTATAAAAGGGATTTAAGCATCGAGGAAC
        : ::::: ::: :::::::::::: ::::: ::: ::::::::: ::::::: :: :::: 
_       AGCAGACCAAGCTGTTTTTGGAAGGAATGCATTACAAAAGGGATCTAAGCATTGAAGAAC
HRIN2-2B1

              600       610       620       630       640       650
_       AGTCAGAATGTACTCAGGACTTTTACCAAAATGTGGCTGAAAGAATGCAGACCCGTGGGA
        ::::::: ::: ::::::: :: ::::: :::::::: ::::: ::::: :: :::::::
_       AGTCAGAGTGTGCTCAGGATTTCTACCACAATGTGGCCGAAAGGATGCAAACTCGTGGGA

              660       670       680       690       700       710
_       AAGTGCCTCCAGAGAAAGTGGAGAAGATAATGGATCAGATCGAAAAGCACATCATGACGC
        :::::::::::::: : ::: :::::::::::::::::::: :::::: :::::::::: :
_       AAGTGCCTCCAGAAAGAGTCGAGAAGATAATGGATCAGATTGAAAAGTACATCATGACTC

              720       730       740       750       760       770
_       GTCTCTATAAATTTGTGTTCTGCCCAGAGACTACTGATGATGAGAAGAAAGATCTCGCCA
        :::::::::::: ::: ::::: ::::: :::::::::::::::::::::::::::: ::::
_       GTCTCTATAAATATGTATTCTGTCCAGAAACTACTGATGATGAGAAGAAAGATCTTGCCA

              780       790       800       810       820       830
_       TTCAAAAAAGAATCAGGGCCCTGCACTGGGTAACGCCTCAGATGCTCTGTGTCCCTGTCA
        ::::::: :::::::: ::::::: :::::: :::::::::::::: :::::::::: :
_       TTCAAAAGAGAATCAGAGCCCTGCGCTGGGTTACGCCTCAGATGCTGTGTGTCCCTGTTA

              840       850       860       870       880       890
_       ATGAGGAAATCCCTGAAGTGTCCGACATGGTGGTGAAAGCGATCACAGACATCATTGAGA
        :::: :: ::::: :::::::: :: :::::::::: :::::::::: ::::::::: :
_       ATGAAGACATCCCAGAAGTGTCTGATATGGTGGTGAAGGCGATCACAGATATCATTGAAA

              900       910       920       930       940       950
_       TGGACTCAAAGCGTGTGCCTCGGGACAAGCTGGCCTGCATCACCAGGTGCAGCAAGCACA
        :::: :: :::::::::::: :::::::::::::::::::: ::::::::::::::
_       TGGATTCCAAGCGTGTGCCTCGAGACAAGCTGGCCTGCATCACCAAGTGCAGCAAGCACA
```

FIG. 11A

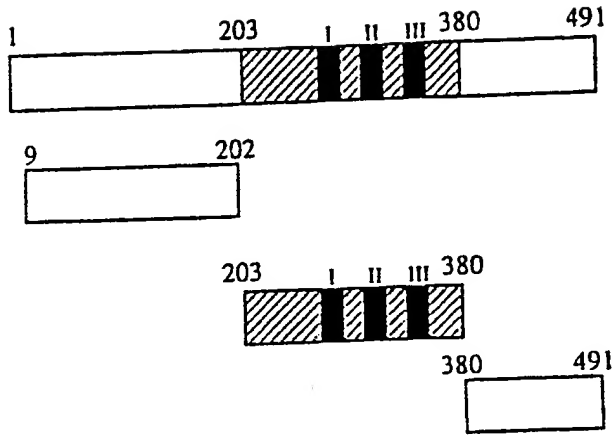
Murine Rin2/ Human HRIN2-2B1 clone (continued)

```
                440        450        460        470        480        490

Murine Rin2 960        970        980        990       1000       1010
TCTTCAATGCCATCAAGATCACCAAGAATGAGCCAGCCTCTGCCGATGACTTCCTGCCCA
:::::.:::::::::::::::::.:::.::::::::: :: ::::: ::::::.::::: :::::
TCTTCNATGCCATCAAGATCNCCANGAATGAGCCGGCGTCTGCGGATGANTTCCTCCCCA
HRIN2-2B1

          1020       1030       1040       1050       1060       1070
CCCTGATCTACATCGTCCTGAAGGGCAACCCCCCTCGCCTGCAGTCCAACATCCAGTACA
:::: :::::: : ::  ::::.:::: :::::::.: ::: :: :: :: ::::: :: :
CCCTCATCTACTT-GTTTTGAANGGCA-CCCCCCNC-CCTTCA-TCTAATATCCA-TATA

          1080       1090       1100       1110       1120       1130
TCACTCGCTTCTGCAACCCCA-GCCGGCTCATGACGGGCGAGGATGGCTACTACTTCACC
:::: ::::::::: : :::: :::: :: :: :: :: .: ::.::::::::: ::: ::
TCACGCGCTTCTGCCATCCCAAGCCGACTGAT-ACTGGANAAGANGGCTACTA-TTC-CC

          1140       1150       1160       1170       1180       1190
AACCTGTGCTGTGCTGTGGCTTTCATTGAGAAATTAGACGCCCAGTCTTTGAATTTAAGT
 : ::::::::::.:::::::::::::: ::: :: :: :::..:::::::::: ::: :
CATCTGTGCTGTGCNGTGGCTTTCATTGAAAAACTAAAC-CCCNNTCTTTGAATCTAATT

          1200       1210       1220       1230       1240       1250
CAGGAGGATTTTGACCGGTACATGTCTGGCCAGACATCCCCCAGGAAGCAGGAGTCTGAG
:.:::: :: ::::: :. :::.: : ::::::.:: : ::.:: .::: :  .:  ::.:
CNGGAAGAATTTGATCNCTACNTTTTTGGCCNGA-ACCCNCCCNGAAACCANAAACTNAA

          1260       1270
AGTTGGCCCCCGGAGGCCTGCTTAGG
: :::: ::::: :.:: ::::::::
AATTGGTCCCCG-ANGCTTGCTTAGG
          790        800
```

FIG. 11B

```
         10         20         30         40
         |          |          |          |
TNCCAANAANATATTCAAGTNTTTTTGGGGGGGCNAAATT 40
CACCCCNTNCGGTTAAAGGAACCAATTGATNTTTTTTTAA 80
CAAAACCNGGGAGTTTTTCCTTNCGGGGGGGGTAAATANG 120
GGGNAACCCAAAGATTNTTGCNATTCAATGCACAGGNGGG 160
ATGTNAACTAAAAACGGAGTTAAATATTTTAGGGGGGTTG 200
        210        220        230        240
         |          |          |          |
ACAGCAACCTGCATNGTAGAACCTTTTTTTTTNTTTCNGNG 240
GACNTTNTATAACNTAAATATACCATTGATGATTTTNTTC 280
CATTCAGTGACATCCACAGATTANGCAGCTATACTTGTGA 320
AATCGTGCATGAGGCCCCCAGGGCACCGTTTTAGAACAAC 360
GTCACTTCACACAGGCAGGTGAGAAAGGTTCTCTTGCTTT 400
        410        420        430        440
         |          |          |          |
TCCAGTATCTTCCTAAGGATGGAGCCCAAAATTGCAGAGC 440
AGTAACTTTGGAATAAAACCAGGGTGGGTATAAAACTTCT 480
TATTCTTAAATTTACATATAAGATCTATTAAGCTTGACAC 520
ATCTGTGTCATCACGCACTGAAGACAGGAAGCAGTTCACT 560
GAGTCAGCTGGTTCCCAAGCTCGCACAGAAGGTGATAAGT 600
        610        620        630        640
         |          |          |          |
TACTATCAAATGCCAGTGAGAATCTTCTTATAGAATAACC 640
TGGGCCCAAGTGATTTTAGTACAAAACTTGCCCTTCTTTG 680
GTTTAATTTTCTATGTGCTTTTAGGTGTGAATCCAGATAT 720
GCGGTCTTAATTCCTTTGGAAATACACAGTTCGTTTAGTT 760
ACTGTACACTCTGTTTGTTCAATAAACTGCATATCAACTT 800
        810        820        830        840
         |          |          |          |
CCAAAAAAAAAAAAAAAAAAA 822
```

FIG. 11C




pACT-Rabex-5-FL

pACT-Rabex-5-D

pACT-Rabex-5-B

pACT-Rabex-5-C

| DNA-binding hybrid | Activation domain hybrid | Colony color | Relative β-Gal units |
|---|---|---|---|
| pAS2-mHRas | pACT-Rabex-5-FL | Blue | 229 |
| pAS2-mHRas | pACT-Rabex-5-D | Blue | 14 |
| pAS2-mHRas | pACT-Rabex-5-B | Blue | 24 |
| pAS2-mHRas | pACT-Rabex-5-C | Blue | 138 |
| pAS2-yRas2p | pACT-Rabex-5-FL | White | 6 |
| pAS2-yRas2p | pACT-Rabex-5-D | White | 4 |
| pAS2-yRas2p | pACT-Rabex-5-B | White | <1 |
| pAS2-yRas2p | pACT-Rabex-5-C | White | <1 |

FIG. 12

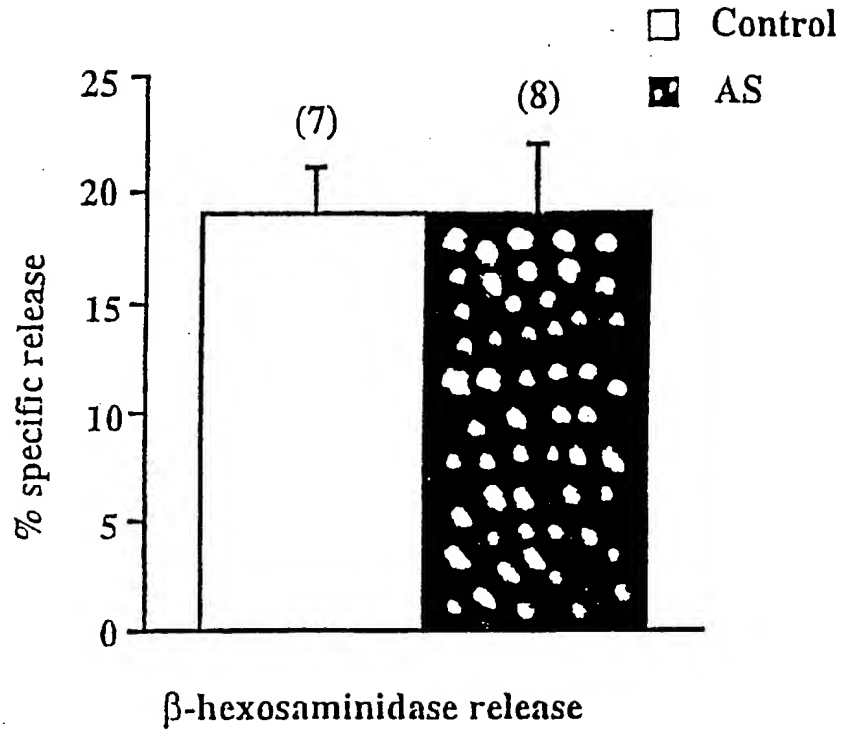



FIG. 13A

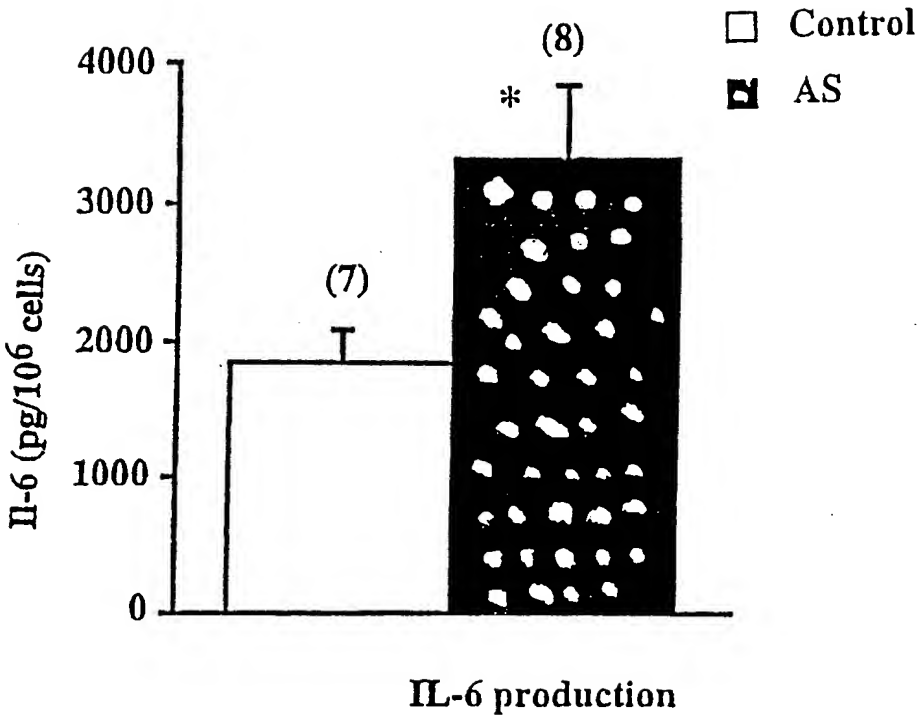



FIG.13B

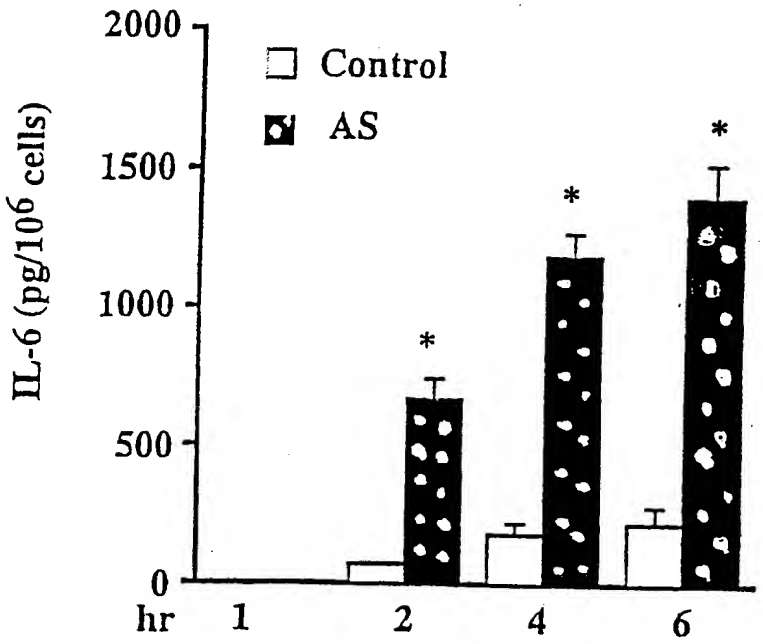



FIG. 13C

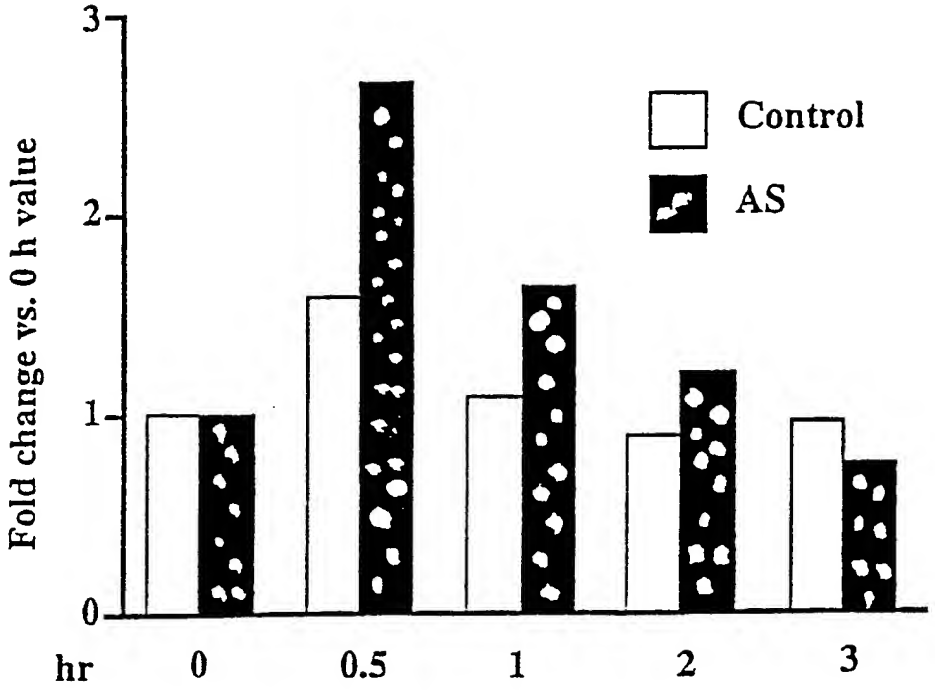

Fold change vs. 0 h value
3
2
1
0
Control
AS
hr
0
0.5
1
2
3

FIG. 14

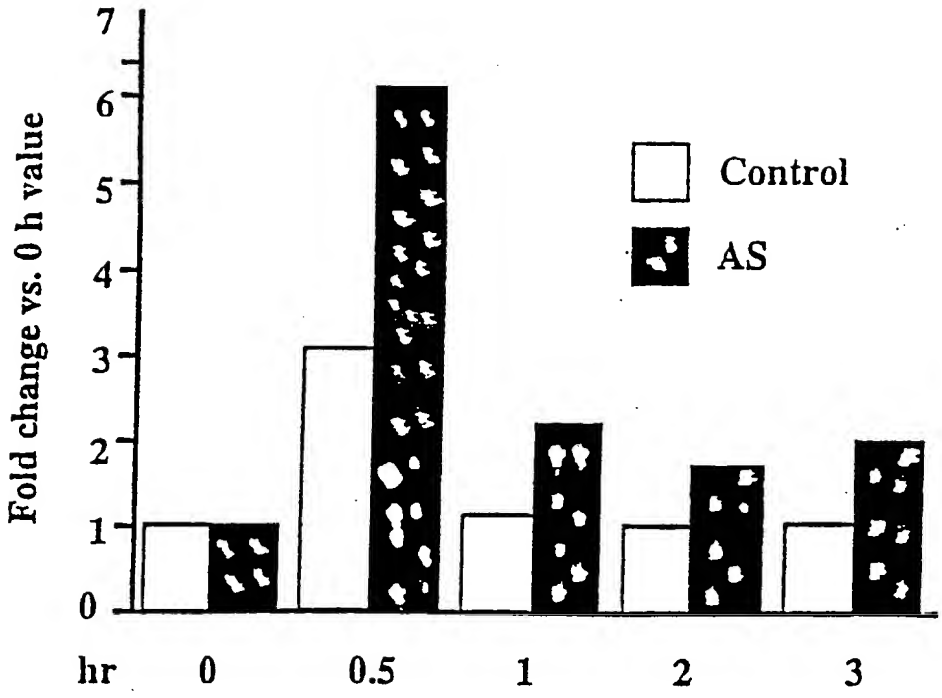

Fold change vs. 0 h value
7
6
5
4
3
2
1
0
Control
AS
hr
0
0.5
1
2
3

FIG. 15

# INTERNATIONAL SEARCH REPORT

Int tional Application No
PCT/US 98/19056

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 C07K16 C12Q1/68 G01N33/50
G01N33/53

According to International Patent Classification (IPC) or to both national class nd IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classific mbols)
IPC 6 C12N C07K C12Q G01N

Documentation searched other than minimum documentation to the extent tha documents are included in the fields searched

Electronic data base consulted during the international sea name of data and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the r- nt passages | Relevant to claim No. |
|---|---|---|
| X | EP 0 666 314 A (COLD SPRING HARBOR LAB)<br>9 August 1995<br>SEQ ID NOS: 13-16<br>see page 12, line 53 - line 56<br>--- | 9,10,12,<br>15,16 |
| X | J. COLICELLI ET AL.: "Expressior of three mammalian cDNAs that interfere with RAS function in Saccharomyces cerevisiae"<br>PROC. NATL. ACAD. SCI.,<br>vol. 88, April 1991, pages 2913-2917,<br>XP002087363<br>NATL. ACAD. SCI.,WASHINGTON,DC,US;<br>see the whole document<br>---<br>-/-- | 9,10,12,<br>15,16 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 10 December 1998 | 28/12/1998 |

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,

Authorized officer
HORNIG H

# INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/US 98/19056

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category | Citation of document, with indication.where appropnate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | L. HAN ET AL.: "Protein binding and signalling properties of RIN1 suggest a unique effector function" PROC. NATL. ACAD. SCI., vol. 94, May 1997, pages 4954-4959, XP002087364 NATL. ACAD. SCI.,WASHINGTON,DC,US; cited in the application see the whole document | 9,10,12, 15,16 |
| X | C.G. BURD ET AL. : "A yeast protein related to a mammalian ras-binding protein, Vps9p, is required for localization of vacuolar proteins" MOL. CELL. BIOL., vol. 16, no. 5, May 1996, pages 2369-2377, XP002087365 ASM WASHINGTON, DC,US cited in the application see the whole document | 9,10,12, 15,16 |
| A | H. STENMARK ET AL.: "Rabaptin-5 is a direct effector of the small GTPase Rab5 in endocytic mmebrane fusion" CELL, vol. 83, 3 November 1995, pages 423-432, XP002087366 CELL PRESS,CAMBRIDGE,MA,US; see the whole document | 1-23 |
| P,X | HORIUCHI H ET AL: "A novel Rab5 GDP/GTP exchange factor complexed to Rabaptin-5 links nucleotide exchange to effector recruitment and function." CELL, (1997 SEP 19) 90 (6) 1149-59. JOURNAL CODE: CQ4. ISSN: 0092-8674., XP002087367 see the whole document | 1-8,11, 13,14, 17-21 |

# INTERNATIONAL SEARCH REPORT

ernational application No.
PCT/US 98/ 19056

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [ ] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. [X] Claims Nos.: 24-54
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
Please see Further Information sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

information on patent family members

Inte ional Application No
PCT/US 98/19056

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| EP 0666314 A | 09-08-1995 | CA 2080920 A | 21-10-1991 |
| | | EP 0537173 A | 21-04-1993 |
| | | WO 9116457 A | 31-10-1991 |
| | | US 5527896 A | 18-06-1996 |

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Claims Nos.: 24-54

Claims 24 relating to agents which alter Rin2 activity relating to claims 25-57 could not be searched as the subject matters were not sufficiently disclosed.

Furthermore Claims 9,10,12,15,16, make no sense SEQ ID NOS. 12,13,14,15 are not nucleotide sequences but amino acids